Anticipatory effects of intonation: Eye movements during instructed visual search

Kiwako Ito *, Shari R. Speer

Department of Linguistics, The Ohio State University, Linguistics College of Humanities, Columbus, OH, USA

Abstract

Three eye-tracking experiments investigated the role of pitch accents during online discourse comprehension. Participants faced a grid with ornaments, and followed prerecorded instructions such as ''Next, hang the blue ball'' to decorate holiday trees. Experiment 1 demonstrated a processing advantage for felicitous as compared to infelicitous uses of L + H* on the adjective noun pair (e.g., *blue ball* followed by *GREEN ball* vs. *green BALL*). Experiment 2 confirmed that L + H* on a contrastive adjective led to 'anticipatory' fixations, and demonstrated a ''garden path'' effect for infelicitous L+H* in sequences with no discourse contrast (e.g., *blue angel* followed by *GREEN ball* resulted in erroneous fixations to the cell of angels). Experiment 3 examined listeners' sensitivity to coherence between pitch accents assigned to discourse markers such as '*And then*,' and those assigned to the target object noun phrase.
© 2008 Published by Elsevier Inc.

Keywords: Intonation; Pitch accent; Contrastive discourse status; Anticipatory eye movements; Discourse marker

When speakers communicate with one another in the course of everyday conversation, they convey a great deal of information beyond their chosen words and sentences. Facial expressions and gestures convey aspects of messages that would be unavailable in the absence of face-to-face articulation. Even when interlocutors are not visually co-present for conversation, listeners have access to a great deal of information beyond the spoken strings of consonants and vowels, including variation in rhythm, melody, tempo, loudness, tenseness, and tone of voice. Intonation provides an organizational structure for speech, and can covey simultaneously a speaker's attitude, utterance purpose, and the relative importance of particular words or phrases. Producing and respond-

ing to various intonation patterns that are specific to a given linguistic community is an automatic, sophisticated, and highly general cognitive skill. For example, most people are sensitive to dialectal differences in intonation (e.g., consider how you would identify the speaker from either Sydney, Boston, or Mumbai by listening to the melody of ''Did he actually show up?''). This sensitivity to dialectal distinctions suggests that the members of a language-speaking community have come to a subconscious consensus on the conventionalized use of intonation patterns.

Consistent with this notion, recent work on intonation strives toward standard descriptions of tunes and their meanings in various languages (e.g., English: Pierrehumbert & Hirschberg (1990); German: Fery (1993), Kohler (2005), Grice, Baumann, & Benzmüller (2005); Italian: D'Imperio (2000); Grice, D'Imperio, Savino, & Avesani (2005); Japanese: Pierrehumbert &

---
* Corresponding author.
  E-mail address: Ito@ling.osu.edu (K. Ito).

Beckman (1988); Venditti (1997, 2005); Spanish: Face (2001); Prieto, van Santen, & Hirschberg (1995). However, this work also reveals a wide range of intonation variation across speakers within the same community, and even within the same experimental setting (cf. Schafer, Speer, & Warren, 2003). Given this diversity of form, how reliably can intonation cue pragmatics during speech comprehension? If intonation is indeed a powerful auditory cue that guides the interpretation of incoming utterances, models of speech comprehension must properly describe the mechanism of such processing. The research presented here uses eye-movement monitoring to investigate whether and when listeners make use of intonational cues to guide visual search through a set of real-world objects as they follow spoken instructions to decorate a series of holiday trees.

The role of pitch accent in American English

According to autosegmental accounts of prosody (e.g., Bruce, 1977; Goldsmith, 1976, 1990; Pierrehumbert, 1980), local excursions of tonal prominence in speech are called *pitch accents*. In English, pitch accents are aligned with the lexically stressed syllable of a word, and are generally accompanied by increased amplitude and duration, and hyper-articulation of the segments as compared to unaccented syllables (Beckman, 1996; Ladd, 1996). Within the framework of a widely used tone annotation system, ToBI (Tones and Break Indices: Beckman & Ayers, 1997), there are five pitch accent types in American English (H*, L*, L* + H, L + H*, and H + !H*: (*) indicates the association or phonetic alignment between the tone and the stressed syllable; (!) indicates the 'downstep,' or contextually triggered lowering of a tone).

These pitch accents are claimed to convey different pragmatic meanings for the utterances in which they appear. Pierrehumbert & Hirschberg (1990) introduced the idea of *compositionality of tonal meaning*, where meanings conveyed by pitch accent, phrasal accent, and boundary tone combine to contribute to the overall meaning of an utterance in discourse. For example, H* accented items are added to the interlocutors' mutual belief space when they appear in either a declarative sentence with final L− (phrasal) and L% (boundary) tones, or in a question with final H- and H% tones. In contrast, the L* accent can also signal the salience of an accented item, but in this case the item is assumed to be already part of the hearer's mutual beliefs. L + H* evokes contrast in the discourse, signaling that "the accented item -and not some alternative related item-should be mutually believed (p. 296)." This is the tune that may be used for the capitalized word in an utterance such as "I made a reservation for FIFTEEN, not fifty!".

Pitch accent in language processing

Given the distribution of pitch accents present in American English, how do speakers and listeners make use of such cues during language processing? We focus here on the contribution of 'emphatic' or 'contrastive' pitch accent to discourse processing. Recent studies of language production have established that speakers' use of pitch accent is tightly tied to the information structure of their utterance as it relates to the discourse structure in which it is embedded. Early work in the area characterized this correspondence as a tendency to accent 'new' information and to refrain from accenting 'old' information (Bolinger, 1961, 1986; Chafe, 1974; Cruttenden, 1986; Halliday, 1967; cf. Chafe, 1976). However, advancements in the specification of models of information structure as well as in the characterization of lexical and sentence level stress and accent have revealed a relationship between accent and discourse status that is considerably more complex. More recent work shows that there is not a strict parallel association between accentuation/deaccentuation of words and their 'new' or 'old' status in a discourse (Bard & Aylett, 1999; Hirschberg, 1993; Ito, Speer, & Beckman, 2003; Nakatani, 1993, 1997; Terken & Hirschberg, 1994). Instead, the correspondence between information structure and speakers' intonational expression of prominences requires the specification of not just the presence/absence of accent on the words that refer to discourse entities, but the particular accent type assigned (Beckman & Pierrehumbert, 1986; Pierrehumbert, 1980; Pierrehumbert & Hirschberg, 1990).

A parallel stream of language comprehension studies have clearly established that the way a referring word is pronounced can influence a listener's speed and accuracy in recovering the speaker's intended referent (see Cutler, Dahan, & Donselaar, 1997, for review). Numerous psycholinguistic studies show higher acceptability ratings and faster comprehension times for listeners when word-level intonation felicitously marks the discourse status of words than when it does not (Bock & Mazzella, 1983; Birch & Clifton, 1995; Needham, 1990; Nooteboom & Kruyt, 1987; Terken & Hirschberg, 1994; Terken & Nooteboom, 1987). In general, felicitously accented words are recognized faster, remembered better, or perceived as more prominent and intelligible than words without accent (Bard, Sotillo, Anderson, Doherty-Sneddon, & Newlands, 1995; Krahmer & Swerts, 2001; Sedivy, Tanenhaus, Spivey-Knowlton, Eberhard, & Carlson, 1995). Deaccentuation or melodic attenuation of previously mentioned information has also been shown to facilitate utterance comprehension (Needham, 1990; Nooteboom & Kruyt, 1987; Terken & Hirschberg, 1994; Terken & Nooteboom, 1987). Although these studies demonstrate that accent plays an important role in helping listeners determine referents, little has been done

to investigate how rapidly pitch accent information is processed and the time course by which it is utilized during speech comprehension. The present experiments investigate listeners' processing of intonational patterns using eye-tracking methodology, which enables the analysis of continuous on-line responses to various intonational patterns as they unfold over time.

The merits of head mounted eye-tracking methodology with naturalistic interactive tasks for the investigation of real-time spoken language processing are well-established (Allopenna, Magnuson, & Tanenhaus, 1988; Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Dahan, Magnuson, & Tanenhaus, 2001; Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Dahan, Tanenhaus, & Chambers, 2002; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995, 1996, see Trueswell & Tanenhaus, 2005 for additional studies and discussion of the advantages of this method). The primary advantages of this method for the research presented here are (1) The head mounted eye-tracking system provides correction for head position and rotation, allowing participants to move and talk during listening data collection. (2) Eye movement monitoring is a continuous, non-intrusive, implicit measure of processing difficulty. Because it allows observation throughout the time course of listeners' perception and interpretation of intonational patterns, we can compare the relative appropriateness of particular intonation patterns during the course of a naturalistic conversational exchange. (3) Established linking hypotheses between spoken word recognition and eye-movements (e.g., Tanenhaus & Trueswell, 1995) allow the assumption that listeners' attention shifts in response to words relevant according to the available spoken context and visual scene, resulting in the non-strategic planning of eye movements to a mentioned object (here, a holiday ornament that must be moved from a grid to a tree). Despite such merits, the experimental paradigm is not free of potential problems. As Trueswell & Tanenhaus (2005) point out, limiting the complexity of visual world or the interactive task with a simple set of materials or actions may not yield the responses that represent normal speech comprehension processing or discourse behavior beyond the experimental environment. This *closed-set* problem is tightly related to unwelcome task-specific strategies. We attempted to avoid these problems by designing our experimental task and visual world to be natural yet fairly complex. Ornament grids displayed more than 40 objects, so that listeners were engaged in genuine visual search rather than simple matching of spoken words to a small set of objects whose locations and identities could potentially be simultaneously held in memory throughout a trial. Participants needed to focus not only on the action of selecting the correct ornament from the board, but also on placing each small ornament

in the correct location on the tree. (See General Discussion for our further thoughts on visual complexity and responses to prosodic prominence.)

Effect of contrastive L + H* during visual search

In a study particularly relevant to the research presented here, Dahan et al. (2002) demonstrated the immediate effect of accentual prominence on eye movements reflecting spoken language comprehension. The effect of accent on reference resolution was shown by using two cohort items such as *candle* and *candy*, which were presented on a computer monitor with phonetically unrelated items such as *necklace* and *pear*, and geometric landmarks such as triangle and square. On each trial, participants heard a paired instructions such as "Put the *candy* above the triangle. Now put the *CANDLE* above the square.", and were told to move the objects accordingly. The target in the second sentence was either repeated from the first sentence (e.g., *candle–candle*) or switched to the cohort competitor (e.g., *candy–candle*), and it was presented either with or without accentual prominence (e.g., *CANDLE* vs. *candle*). Participants' eye movements to the target object mentioned in the second sentence were monitored. The proportion of fixations to the cohort competitor (e.g., *candy*) was higher when the repeated target carried prominence (e.g., *candle–CANDLE*) than when it did not (e.g., *candle–candle*). In addition, when the target was not repeated from the first to the second instruction (e.g., *candy–candle*), the effect was reversed: fixation proportions to the competitor (*candy*) were higher when the target was not prominent than when it was. Dahan et al. interpret these results as indicating the anaphoric interpretation of non-prominent words and nonanaphoric interpretation of prominent words.

This is partially consistent with the claim by Terken & Nooteboom (1987) that a deaccented entity is processed as *given, and thus* a listener tries to match that entity to an already-activated discourse entity, whereas the interpretation of an accented word is not constrained in such a manner. The results echo previous off-line studies showing that improperly accenting an already-mentioned entity or deaccenting a yet-to-be-mentioned entity increases comprehension time (Bock & Mazzella, 1983; Nooteboom & Kruyt, 1987; Birch & Clifton, 1995). Note, however, that statement-verification tasks show that infelicitous accentuation of *given* (i.e., already-mentioned) items does not slow down comprehension as much as infelicitous deaccentuation of *new* (never-mentioned) items (Birch & Clifton, 1995; Ito, 2002). This may mislead us to conclude that the presence of an accent does not have an immediate impact on the interpretation of incoming speech. The eye fixation pat-

terns shown by Dahan et al. demonstrate that this is not the case. Instead, accent leads to rapid identification of an item not previously mentioned within the interactive environment.

Interestingly, Dahan et al.'s stimuli contained a variety of pitch accent and phrasal tone combinations. According to their ToBI annotation, 14 out of 24 targets in the accented condition had L + H* while the remaining 10 cases had H*. In the deaccented condition, all targets carried H + !H*, but a preceding discourse marker (e.g., on 'Now,' Experiment 1) had either L + H*, L* + H, or H*. In addition, the theme (e.g., *candy*) in the first instruction had L + H* in 26, H* in 20, and L* in 2 of 48 items. The combined phrasal–boundary tones varied between L–L% and L–H%. This variation of accent pattern in the context instruction may have affected the salience and informational status of target or competitor items even before subjects heard the second instruction. More specifically, it is possible that contrastive accents in the context instruction may have contributed to the finding of a very early advantage for items with felicitous accent patterns (shown within 300 ms of target onset). The absence of phonetic control over auditory stimuli is problematic for a comparison across studies (see General Discussion for a close comparison between the present study and Sedivy, Tanenhaus, Chambers, & Carlson (1999), which failed to demonstrate the effect of focus intonation). In the present study, we controlled sentence-level tonal patterns for all trials including fillers. Stimuli were re-recorded until two ToBI annotators agreed on their respective tonal patterns. In order not to make the target instructions sound noticeably contrastive or monotonous, intonation variation was introduced by manipulating the patterns in filler trials (see Materials sections below). Furthermore, we specifically tested the effect of accentual patterns on the discourse marker in the last experiment.

A more important goal of the current study is to explore the additional role of contrastive accent L + H*, in hope of advancing the investigation of a general mechanism of accentual information processing during speech comprehension. As demonstrated by Dahan et al., prominent accent on the noun naming the direct object of the task action 'put' promptly assigns the new/not-previously mentioned or given/already-mentioned status to the target entity. Obviously, the tonal shape of a word plays a role beyond simply signaling its informational status with respect to discourse background. The present study entertains the possibility that contrastive accentual information feeds the discourse *foreground*, projecting pragmatic links between the accented discourse entity and upcoming entities. Compare, for example, what you would expect as continuations of utterance fragments (1) and (2).

(1) KATIE did not win a truck,
     L+H*   H+!H* L − H%
(2) Katie did not win a TRUCK,
     H*   L+H* L − H%

In (1), it is very plausible that the sentence continues with somebody else's name who won a truck (e.g., "...,*LAURA did*)", whereas (2) may continue to mention something other than a truck that Katie actually won (e.g., "...,*she won a MOTORCYCLE*)". Such predictions become available due to the contrastive function of L + H*, which evokes a set of alternatives appropriate for the immediately ensuing discourse (Pierrehumbert & Hirschberg, 1990; for the effect of contrastive accent in negation, see Davidson, 2001). In other words, a contrastive L + H* constrains the set of coherent upcoming discourse entities as well as specifying the informational status of the accented word itself. Consistent with the immediate effect of accent demonstrated by Dahan et al., we predict that anticipatory effects of L + H* will be evident in eye movements when a member of the evoked set is present in the visual field. Furthermore, we pursue the view that accentual cues are evaluated immediately, allowing assignment of referential status to the named discourse entity as its speech segments unfold in time. That is, a particular pragmatic relation is established between the accented word and other background/foreground discourse entities during the recognition of the content of the speech signal. Although the immediacy of accentual processing itself may not be contentious, the question remains as to whether the processing of accentual information is complementary to incremental speech comprehension, or if instead it is a robust independent computation that takes place in parallel with phoneme-based word recognition. Dahan et al. suggest strong influence of parallel prosodic processing on lexical access, but their evidence is rather indirect due to the use of cohorts. In the present study, we manipulated the accentual prominence of a prenominal adjective (Experiments 1 and 2) to test the effect of contrastive accent on the modifier on the access to the upcoming noun. If a prominent accent on a prenominal adjective evokes a contrastive relation between the immediately preceding referent and the about-to-be-mentioned referent, it should constrain the set of nouns that the adjective modifies. Hereafter, we refer to this constraint on the upcoming referent as an "anticipatory effect" of contrastive accent. We hypothesize that this effect of prosodic prominence is independent of the segmental processing necessary for lexical identification, and thus predict that speech comprehension will be disrupted when the segmental information of the noun does not substantiate the anticipated referent (The effect of such

mismatch between the anticipated referent and the incoming noun is directly tested in Experiment 2). The scope of the anticipatory effect in discourse foregrounding is tested by manipulating accentual patterns of discourse markers that connect utterances in Experiment 3.

We now turn to a series of three experiments, where participants decorated holiday trees following instructions that mention target ornaments as combinations of color adjectives and object nouns (e.g., blue ball). Experiment 1 tests whether felicitous use of L + H* (e.g., "*First, hang the green ball*" → "*Now, hang the BLUE ball*") facilitates fixations to a target object as compared to infelicitous use (e.g., "*First, hang the green ball*" → "*Now, hang the blue BALL*"). Experiment 2 confirms the facilitatory effect of felicitous L + H* by comparing it against a condition that does not involve L + H* (e.g., "*First, hang the green ball*" → "*Now, hang the blue ball*"). This experiment also examines whether the set-evoking function of L + H* can mislead listeners to fixate on the previously mentioned object type in the presence of the name of a different object (e.g., Will balls be fixated during the second instruction for the sequence "*First, hang the green ball*" → "*Now, hang the BLUE angel*"?). Finally, Experiment 3 investigates whether the presence of L + H* on temporal-adverbial discourse markers such as *And then*, *After that*, and *And next* have additional effects on the processing of the upcoming target.

Experiment 1

The first experiment was designed to test whether felicitous use of a contrastive pitch accent (here, L + H*) is advantageous for listeners as compared to infelicitous use during the holiday tree decoration task. We hypothesized that L + H* on a color adjective evokes a set of alternative color possibilities that could modify the ornament object, while restricting the referent to be named by the upcoming noun to the same type of object that was mentioned in the immediately preceding utterance. According to this view, L + H* accent on the color adjective should not only assign contrastive status to the color itself, but also should increase listeners' expectations that the most recently mentioned target noun will be repeated in the current utterance. We predicted a rapid anticipatory effect of the adjective's intonational prominence on the listener's selection of a candidate noun, which might be detectable even before the noun's segmental information was fully processed for word recognition, consistent with previous demonstrations of the very early use of prosodic information in visual world tasks (Dahan et al., 2002; Snedecker & Trueswell, 2003).

In all experiments in this study, ornaments were sorted by type into the cells of a display grid (e.g., one cell contained all balls, another all angels, etc.). We predicted that fixations to the target ornament cell would be speeded when L + H* on the adjective felicitously marked contrast between two consecutive referents in sequences such as "*Hang the green drum. Now, hang the BLUE drum.*" In contrast, fixations would not be speeded if L + H* expressed contrast in an infelicitous position, e.g., on the object noun instead of the color adjective, as in "*Hang the green drum. Now, hang the blue DRUM*". In such cases, intonation-based anticipation would not be established at the adjective, and despite the prominence on the object noun itself, fixations to the target cell should not be as early as those for the felicitous sequences. Experiment 1 also examined whether felicitous use of contrastive L + H* on the object noun provided a processing advantage to listeners during visual search. We compared sequences such as "*Hang the green drum. Now, hang the green BALL*" to infelicitous use in sequences such as "*Hang the green drum. Now, hang the GREEN ball*". Notice that noun contrast trials differed from adjective contrast trials in that it was possible for the lack of contrastive accent in combination with the repetition of the adjective in the felicitous sequence (green drum → green BALL) to serve as a cue to upcoming noun contrast even before the occurrence of the contrastive L + H* accent. However, because the visual search was among cells organized by object type, such a cue could only serve to foreshadow contrast and could not allow projection of a particular noun candidate (because there were always multiple green ornaments in the display). Thus, we predict no difference between felicitous and infelicitous noun contrast trials until after the phonemic information of the noun becomes available.

Methods

Participants

Thirty-six native speakers of Midwestern American English were recruited at the Ohio State University. They received partial credit toward a course requirement for their participation.

Design and materials

Visual search task. Each participant had four trees to decorate, each from its own grid of ornaments. Grids held 40 to 52 ornaments, and finished decorated trees held a total of 24 (16 target and 8 filler) ornaments. No grid held more than one ornament of a particular type and color combination. Grid layouts were designed to prevent listeners from being able to predict the mention of an upcoming color or ornament type, either by process of elimination as the trials transpired, or by a pattern of repeatedly choosing items of the same name

across trees. All colors and all ornament types were mentioned at least once, and no more than three times during the decoration of a single tree.

Real-world objects. Eight target color adjectives (blue, red, green, orange, gold, silver, brown, and grey) were combined with eight target object nouns (ball, drum, angel, bell, stocking, onion, candy, and egg) to construct noun phrases that described the ornaments. Three additional colors (purple, white, and yellow) were used to paint four filler ornament types (snowman, lightbulb, tree, and star) as well as 'dummy' ornaments that were displayed on the grid but remained unmentioned in instructions. Color shades and ornament shapes or sizes of target objects were varied across trees, so that each color adjective and each object noun could refer to a different and novel discourse entity in each tree. Filler ornament types were varied across trees.

Adjective–noun sequences. Within each tree, construction of contrastive trials required four sequences that repeated the object noun, serving as a context where the color adjective should convey contrastive information (e.g., *green onion → orange onion)*. There were also four sequences that repeated the color adjective, serving as a context where the object noun should convey contrastive information (e.g., *brown ball → brown angel*). Across the four trees, each adjective and each noun appeared in these contrastive sequences one to three times, preventing the association of particular colors or objects with specific pragmatic contexts. Infelicitous contrastive sequences were limited to two noun and two adjective contrast trials per tree, to minimize the possibility that listeners might adopt some task-specific strategy or become insensitive to pitch accent information due to its infelicitous use. The design also included four non-contrastive trial types that were used to examine the effect of noun or adjective repetition. These trials are not discussed here for reasons of space (see Ito & Speer, 2006, for discussion). Appendix A includes the full ornament sequences for the four trees used in Experiment 1.

The instructed visual search task design used here differed from more traditional psycholinguistic experimental designs in several important respects. First, the visual array of objects remained substantially the same across the course of the experiment, changing only 4 times instead of on each new sentence trial (as in Snedecker & Trueswell, 2003), or each paired sentence trial (as in Dahan et al., 2002). This allowed us to measure the use of intonational information by participants engaged in a complex naturalistic task-based search, arguably a more natural setting for language use. Second, more traditional designs could be said not to involve search at all, but instead involve mapping names recovered from the speech signal to a circumscribed display of otherwise unrelated or loosely related objects (such as frogs and flowers, or candles and candies). Participants in these tasks have little need to search, readily fixating items based on similarity to the spoken input, and rarely looking at unrelated items. Here, the task and the larger number of items require search, and as such the test of whether spoken language processing can be immediately influenced by simultaneously available visual input is more rigorous. Third, because in our task the real world referents of words were not renewed from trial to trial, and groups of real world objects had the same name, the lexical items used named object categories in the display (i.e., colors and ornaments) as well as single objects. This manipulation was necessary in order for us to examine the contrastive function of L + H* accents in the context of a set. Fourth and finally, the design is not appropriate for the traditional psycholinguistic statistical analysis that considers language items as a random effect, because the necessity for repetition without predictability in the creation of the conditions made it impractical to use the same words in all conditions for reasons of time. Item analyses assume random sampling of items and are used to reduce the probability of Type I error due to differences between item sets in conditions. They are often used to eliminate the contribution of particular lexical items to sentence processing effects, where factors such as lexical frequency or coherence across a sequence of words may artificially inflate or reduce processing time for a particular item. Here, the task structure introduced items in critical conditions into the discourse structure by visual presence and by mention of their names before they were mentioned as part of a target noun phrase, minimizing the contribution of initial lexical access or coherence to the processing of individual items. Nevertheless, we matched the sets of spoken items for number of syllables and absolute duration within conditions to be critically compared, as these variable have been shown in some studies to have a positive correlation with reaction time, and thus might have influenced time to fixate the target. All words had first-syllable stress. (Mean number of syllables and duration in critically compared conditions: Felicitous: adjective 1.3, 337 ms; noun 1.5, 449 ms; Infelicitous: adjective 1.3, 332 ms, noun 1.6, 469 ms; Adjective contrast: adjective 1.3, 327 ms, noun 1.6, 457 ms; Noun contrast: adjective 1.3, 343 ms, noun 1.6, 461 ms). When items are matched rather than random, and/or when they are selected from a highly constrained set (such as tree ornament names) use of Fl alone is the correct procedure (See Raaihmakers, Schrijnemakers, & Gremmen, 1999, for discussion of the conditions under which items analyses are not appropriate). To allay any concerns about the generality of our effects, we rotated adjectives and nouns into different conditions across Experiments 1–3, thus replicating effects across different sets of word items in conditions.

Spoken materials. The intonation patterns for the audio instructions were determined based on speech production data from a previous experiment using a comparable holiday tree decoration task (Ito et al., 2003; Ito & Speer, 2006). In the previous study, naive speakers gave instructions on how to decorate holiday trees to our confederate, who decorated them. Speakers were not instructed as to what to say, but saw a CRT display with two color photographs on each trial. One showed the next ornament to hang, and the other showed a tree marked at the appropriate location with a text tag naming the ornament (e.g., 'green onion'). This unscripted interactive task allowed us to record spontaneous productions of the intonation patterns that participants typically produced for the adjective–noun ornament names. Results showed that both adjectives and nouns were very likely to be accented when they were mentioned for the first time (over 80% of the time). However, while adjectives were accented with similar frequency regardless of whether they had been already mentioned, nouns were less likely to be accented on subsequent mentions (probability of accentuation dropped to .5 when an already-mentioned noun followed the first mention of an adjective). When adjectives appeared in a contrastive context, they were frequently pronounced with $L + H*$ (46% when they were mentioned for the first time, and 50% for subsequent mentions). $L + H*$ on nouns in contrastive contexts was less frequent (19% first mention, 18% already mentioned).

Auditory stimuli for Experiment 1 were constructed on the basis of these patterns from naive speakers. Items contained a pitch accent on both the first and consecutive mentions of both nouns and adjectives. Adjective noun pairs that did not occur in contrastive context were assigned the sequence most typical for such pairs in our production study, $H* !H*$, where the accent on the noun was 'downstepped' from the preceding adjective's accent. The only exception was when the first mention of a noun followed an already-mentioned adjective, in which case the sequence assigned was $H* H*$ (again, with reference to such sequences in the production study). In contrastive contexts, where an adjective or noun was repeated on the immediately following trial, $L + H*$ was assigned to the contrast word. When a noun was repeated, $L + H*$ on the adjective was followed by

no accent on the noun. When an adjective was repeated, it was assigned a $H*$, and $L + H*$ occurred on the following noun. Table 1 shows the four critical conditions and their associated pitch accent patterns for the instructions used in Experiment 1. (Four additional conditions appeared in the design of the experiment. These conditions tested the effect of initial vs. repeated mention of adjectives and nouns, but did not involve contrast. They are omitted from further discussion here for reasons of space). The crucial comparisons among the four conditions are between the Felicitous and Infelicitous Adjective Contrast trials (Felicitous Adj Contrast vs. Infelicitous Adj Contrast) and between the Felicitous and Infelicitous Noun Contrast trials (Felicitous N Contrast vs. Infelicitous N Contrast). Felicitous and Infelicitous Adj Contrast conditions appeared in "contrast-on-color" discourse context, where the words describing the sequence of objects specified a new color for a subsequent object of the same type (e.g., *green drum → blue drum*). Felicitous and Infelicitous N Contrast conditions appeared in "contrast-on-object" discourse context, where the words describing the sequence of objects specified the same color for a subsequent object of a different type (e.g., *blue onion → blue drum*). Each tree had two trials in each of the four critical contrast conditions.

Instructions were recorded at 22.05 KHz, 16 bit resolution (using SoundEdit 16, Version II, Macromedia) by a trained female phonetician who could maintain a consistent overall pitch range and intonation pattern across items within each condition. The F_0 values were calculated with a 10 ms window using an autocorrelation algorithm (Boersma, 1993) in Praat (Version 4.2.17: Boersma & Weenick, 1992-2004). To validate the accentual patterns present in the stimuli, two independent ToBI transcribers annotated the target instructions in a randomized order. Instructions were re-recorded until both transcribers gave independent annotations indicating that the target adjective and the noun carried the intended intonational patterns. Fig. 1 shows example ToBI transcriptions of the utterances used for the Felicitous and Infelicitous Adj Contrast conditions.

Our ToBI annotators used the pitch accent categories $H*$ and $L + H*$ as specified in the guide to ToBI labeling (Beckman & Ayers, 1997). However, we acknowledge that some scholars consider the phonological distinction

Table 1
Experiment 1: Information status and accent pattern for the critical four conditions

Type of mention	Preceding context	Target instruction with accent specification
Critical conditions		
Felicitous Adj Contrast	... green drum →	BLUE$_{L + H*}$ drum$_{no\ accent}$
Infelicitous Adj Contrast	... green drum →	blue$_{H*}$ DRUM$_{L + H*}$
Felicitous N Contrast	... blue onion →	blue$_{H*}$ DRUM$_{L + H*}$
Infelicitous N Contrast	... blue onion →	BLUE$_{L + H*}$ drum$_{no\ accent}$

Note: Adj, Adjective; N, Noun. → indicates immediate adjacency between the target trial and the preceding trial.
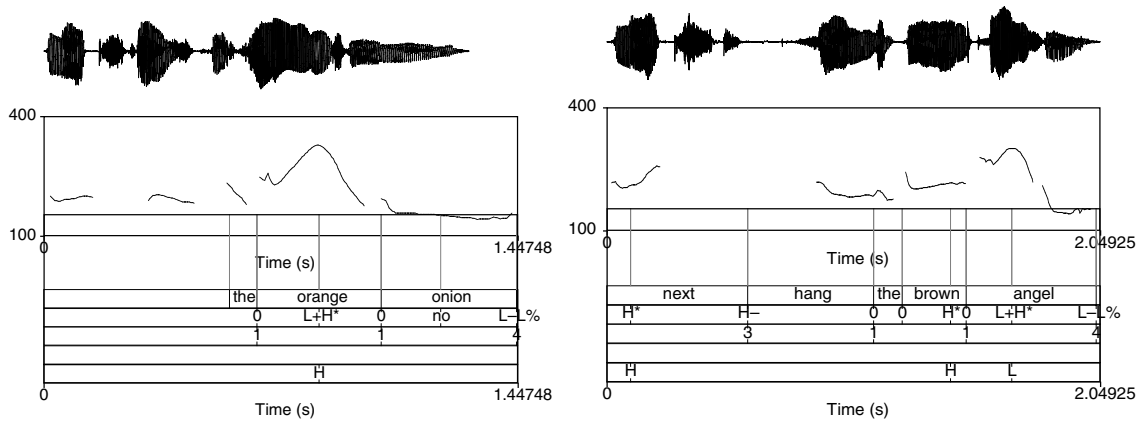
Fig. 1. Example ToBI transcriptions of target noun phrases, (left) *Hang the ORANGE onion* in the Felicitous Adj Contrast condition: and (right) *Next, hang the brown ANGEL* in the Infelicitous Adj Contrast condition.

between H* and L + H* to be continuous rather than categorical. Contradictory results have been found for these accents in comparisons of perceptual rating, classic auditory identification and discrimination tasks (e.g., Bartels & Kingston, 1994; Ladd & Morton, 1997). In Ladd & Morton (1997), listeners rated the degree of emphasis for a set of sentence stimuli (e.g., He's Iranian) created by modifying the F_0 peak range with step sizes 6, 8, l0, and 16 Hz. Emphasis judgment ratings showed a gradual increase with increasing F_0 augmentation, suggesting continuous perception. However, forced-choice identification between "everyday occurrence" and "unusual experience" interpretations of stimuli from the same continua showed the classic *S*-shape response curve, with a sharp increase for the "unusual experience" response in the mid F_0 range, suggesting a categorical boundary between normal and emphatic accents. Results of a same/different discrimination task using paired stimuli from the same series failed to show a peak for the correct "different" responses in the F_0 mid range, which would have indicated a boundary between the two categories. Ladd and Morton argued that these results suggest that pitch range changes are *categorically interpreted*, though they may not be *categorically perceived* (Ladd & Morton, 1997, p. 339). Although further research is needed to determine whether there is a 'true' categorical boundary between H* and L + H* (like the one between phonemes /p/ and /b/), these results do clearly indicate that pitch accent height has a substantial impact on the interpretation of spoken emphasis, with listeners interpreting accents with relatively higher F_0 as more emphatic. In the experiments presented here, we distinguished H* from L + H* on the basis of maximum F_0 in the accented syllable and auditory discrimination by ToBI annotators, choosing unambiguous tokens for use as stimuli in the two categories.

The ToBI-annotated materials for Experiment 1 were submitted to phonetic analysis. Duration values were obtained for the verb, the article *the*, the adjective, and the noun. For accented items, peak F_0 values were measured for H tones. Table 2 shows the mean duration and F_0 values of the target stimuli. Note that words with L + H* clearly differentiated with higher F_0 peaks and longer duration as compared to words with H*.

Procedure

Participants were seated in front of an ornament grid supported by a drafting table with a surface incline of 35 degrees. A small holiday tree was located to the left side of the table on a rotating stand. Location and timing of eye movements were monitored with an *ASL* E5000 head mounted eyetracker with eye-head integration. Eye position was calibrated initially and whenever needed to prevent track loss throughout the experiment.

Table 2
Experiment 1: Mean duration and F_0 values of the target stimuli

Condition	Verb	*the*	Adj			Noun		
	(ms)	(ms)	(ms)	accent	F_0 (Hz)	(ms)	accent	F_0 (Hz)
Felicitous Adj Contrast	343	95	347	L + H*	304	401	—	—
Infelicitous Adj Contrast	337	113	307	H*	207	513	L + H*	291
Felicitous N Contrast	318	133	328	H*	205	498	L + H*	298
Infelicitous N Contrast	348	106	358	L + H*	295	425	—	—

Participants wore light plastic headgear supporting a 60 Hz eye-camera and a small magnetic receiver, which signaled head position to the system. A 60 Hz stationary scene camera was mounted on the ceiling behind the participant, providing a view of the ornament grid. The experimenter sat behind the drafting table, operating a laptop computer running *E-prime*, Version 1.0 (Psychology Software Tools, Inc.). Individual trial onsets were initiated by the experimenter via a button box connected to the laptop. The onset and offset of eye position data collection was synchronized with the onset and offset of the critical auditory phrases in each instruction by an *E-prime* command sequence that played the sound and initiated data recording via a serial connection to the *ASL Control Unit.*

Participants were told that for each trial, they would listen to an instruction, pick the specified ornament from the grid, hang it on the tree at the instructed location, and then face back to the board and say "O.K." so that the experimenter could play the next instruction. Instructions were played through a set of speakers placed on the experimenter's desk, facing the participant. The layout of experimental scene is depicted in Fig. 2.

Each of the four experimental grids had 11 cells, each containing 3–5 ornaments. For each tree, the 16 critical target ornaments (2 in each condition) were placed in the 8 cells surrounding the center cell. Filler ornament sets occupied the center cell and two cells located on the far left and right sides of the grid. Throughout the trials, at least one ornament remained in each grid cell, so that the predictability of color and ornament terms did not change over the course of tree decoration. Fig. 3 shows a photograph of a complete grid at the beginning of a tree sequence.

Results and discussion

Three participants requested that one instruction be replayed, when they did not remember the direction of decoration (e.g., "Moving to the right,"). Since they all fixated the correct cell before reaching the target ornament during the first presentation of instruction, eye movement data during the replay were discarded. Two more trials were discarded from two other participants because the target ornament was missing from the grid due to experimenter error. Across the critical conditions, there was only one missing trial (Infelicitous Adj Contrast Condition: 1/288 observations missing).

Saccades and fixations to target cells were automatically generated using *ASL* E5000 software. Fig. 4 shows the proportions of fixations to the target cells averaged across participants for the four critical conditions, coding fixation onset times from the onset of the saccade (Altmann & Kamide, 2004). The eye movements were recorded at 60 Hz, and thus each sampling point on this figure indicates how likely the fixation was within the target cell at approximately every 17 ms. Proportions
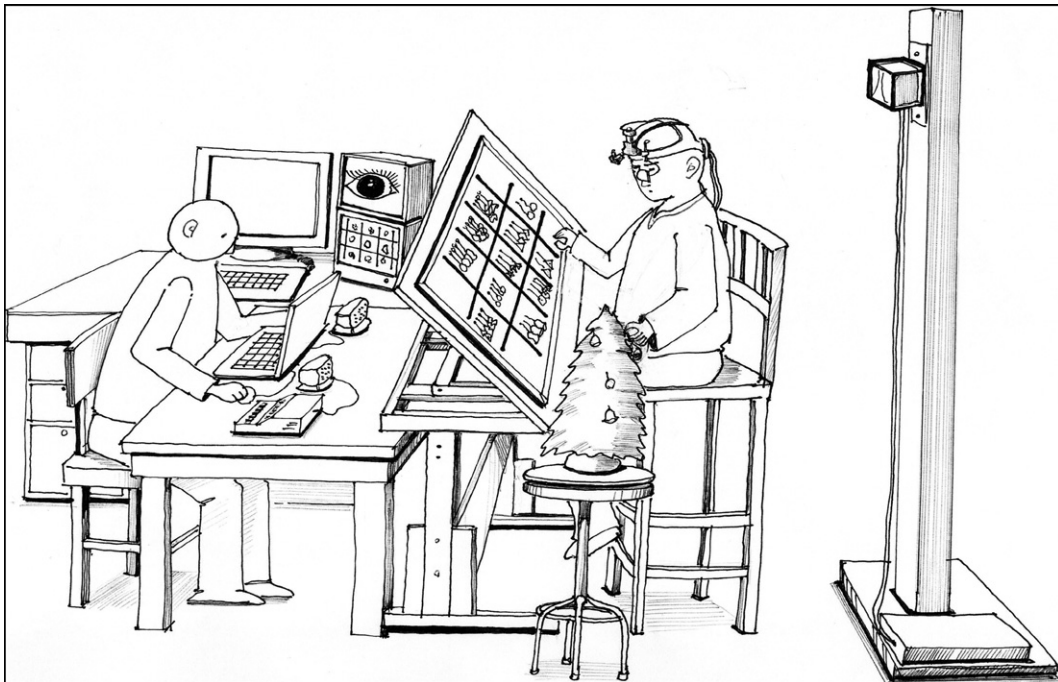

Fig. 2. Experimental setup for holiday tree Experiments 1–3.

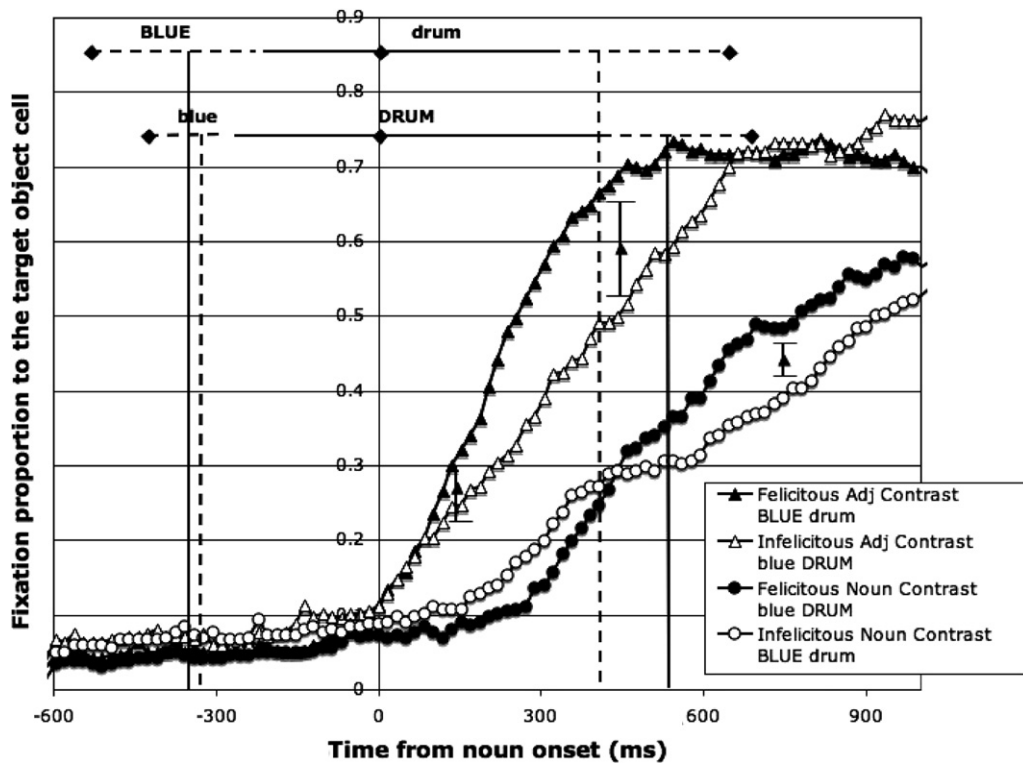Fig. 3. An example image of the ornament board.

Fig. 4. Experiment 1: Proportion of fixations to the target object cell, aligned from noun onset in the four critical conditions.

give the number of fixations to the target cell divided by the number of possible fixations (36 participants × 8 trials per condition, 288). The fixation proportions are shown aligned from the onset of the object noun, displaying the fixations during the prenominal adjectives with negative time values. This alignment strategy was adopted for two reasons. First, the object noun was the critical word that singled out the referent to be mapped onto the real world object. Thus, aligning fixation proportions from the onset of the noun allows direct comparison of eye movement timing between our data and those of previous studies such as Allopen-

na et al. (1988) & Dahan et al. (2002). Second, backward alignment of the data allows the examination of the anticipatory effect of accent toward the end of adjective without risking the loss of data. (Note that double-alignment of data at the onset of adjective and again at the onset of the noun using the average duration of words (as in Snedeker & Trueswell, 2003) would ignore data available from the final portions of words that had longer-than-average durations.) In Fig. 4, the mean duration of adjectives and nouns is indicated by a vertical line for each of the two accentual patterns. Horizontal lines indicate the range of word duration, for each accentual pattern, with solid lines indicating the shortest item duration, and dotted lines extending to the longest item duration. The floating vertical lines indicate the 99% confidence intervals for the time regions where the mean difference reached statistical significance for each critical comparison. Values of the confidence intervals were back-transformed from arcsine-transformed proportions that served as the dependent measures in the statistical analyses.

It is important to note that an alternative method to calculate fixation proportion has been proposed to normalize the durational differences across spoken items (Altmann and Kamide, 2004). According to this method, fixation proportion signifies how often subject fixated the target during each word. This strategy has been adopted by other researchers to analyze the time course of referential resolution in the presence of syntactic ambiguity (Engelhardt et al., 2006). Although this method is appropriate to examine at which word the effect of sentence context or syntactic structure appears during sentence processing, it does not allow detection of the point *within* the critical word where the effect of prosodic cue on word recognition starts appearing. Since it is of our particular interest to identify the point where the effect of contrastive accent appears within the noun phrase, we calculated fixation proportions for each time point from the onset of the noun.

In order to examine the time course of prosodic effects on visual search, mean arcsine transformed fixation proportions for each 300 ms window before and after the onset of the noun were submitted to five (2 backward and 3 forward time windows) 2 (discourse contrast) × 2 (L + H* accent location) repeated measures ANOVAs. This size of time widow was chosen to examine (1) whether the effect of accentual manipulation for the adjective appears during the adjective itself, and (2) whether the accentual manipulation for the adjective affects the fixation patterns before the segmental information of the noun is used for launching the eyes. Here, we assume that saccade planning in a visual search task with a field as complex as our experiment would require 200–300 ms after initial segmental processing (Viviani, 1990). Table 3 summarizes the results of critical pairwise comparisons. The relevant confidence

intervals were calculated with the means by participants. Item analyses are not provided (see discussion in Experiment 1 method section above).

Fig. 4 shows the effect of felicitous L + H* on the prenominal adjective for Felicitous and Infelicitous Adjective Contrast conditions where the ornament type was immediately repeated and thus the color adjective conveyed contrastive information (e.g., *green drum* → *blue drum*). The data show more early fixations to the target cells when L + H* felicitously marked the adjective (*BLUE drum*), expressing its contrastive status, than when L + H* was infelicitously placed on the object noun (*blue DRUM*). Note that the fixation proportion rises from the beginning of the noun in both conditions. However, the initial rise is clearly steeper in the Felicitous Adjective Contrast condition [L + H* no-accent] than the Infelicitous Adjective Contrast condition [H* L + H*]. The two lines diverge at around 100 ms into the noun, with no overlap until after 650 ms. We assume that planning and execution of a saccade requires at least 150–200 ms for simple visual search (Fischer, 1992; Matin, Shao, & Boff, 1993; Saslow, 1967), and, as we have mentioned, the visual search involved in our task is likely to be more complex. Thus, the early more frequent fixations to the target cell in Felicitous Adjective Contrast condition could not have been driven solely by the phoneme-based recognition of the object noun. Instead, we claim that fixation to the target cell was facilitated by the preceding adjective's contrastive accent, which restricted the candidates for the upcoming referent of the noun to the set of ornaments that were the same type as those in the previous trial.

One concern for this interpretation of our results, however, is that the adjectives with felicitous L + H* were on average 40 ms longer than those with infelicitous H*s. Thus the accentual prominence may have simply provided a durational advantage, with longer adjectives serving as additional opportunities for planning saccades to the target. In order to eliminate this possibility, we examined the correlation between adjective duration and the first fixation latencies from the onset of the noun in each condition finding no significant relationship ($r = -.165$; .004; .149; .154 for Felicitous Adjective Contrast, Infelicitous Adjective Contrast, Felicitous Noun Contrast, and Infelicitous Noun Contrast, respectively). Therefore, chances are slim that participants took advantage of longer adjective durations to plan saccades to the target cells.

Fig. 4 also compares Felicitous and Infelicitous Noun Contrast conditions, where the color adjective was immediately repeated (in *contrast-on-object* discourse context). Unlike the above comparison in contrast-on-adjective context, there is no clear early advantage for the felicitous condition, where the adjective carried a H* accent and L + H* properly marked the contrastive status on the noun. In fact, fixation proportion data

Table 3
Experiment 1: Results comparing fixation proportions ($n = 36$), Subject analysis

Time region (ms)	Mean fixation proportion	Mean fixation proportion	Mean Difference Felicious–Infelicitous	CI	F
Felicitous Adj Contrast vs. Infelicitous Adj Contrast					
−600 to −300	.051	.069	−.018	.041	.804
−300 to 0	.065	.088	−.023	.047	.929
0 to 300	.308	.238	.073	.075	3.886^
300 to 600	.684	.506	.221	.116	27.59**
600 to 900	.766	.724	.062	.109	1.365
Felicitous Noun Contrast vs. Infelicitous Noun Contrast					
−600 to −300	.041	.067	−.026	.041	1.607
−300 to 0	.057	.078	−.021	.046	.855
0 to 300	.086	.120	−.034	.075	.852
300 to 600	.265	.273	−.008	.083	.033
600 to 900	.491	.394	.108	.044	4.031**
Felicitous Adj Contrast [L + H* no-acc] vs. Infelicitous Noun Contrast [L + H* no-acc]					
−600 to −300	.051	.067	−.016	.042	.661
−300 to 0	.065	.078	−.013	.048	.295
0 to 300	.308	.120	.192	.100	27.35**
300 to 600	.684	.273	.460	.116	126.4**
600 to 900	.766	.394	.450	.146	75.62**
Felicitous Noun Contrast [H* L + H*] vs. Infelicitous Adj Contrast [H* L + H*]					
−600 to −300	.041	.069	−.028	.042	1.825
−300 to 0	.057	.088	−.031	.048	1.810
0 to 300	.086	.238	−.153	.100	17.49**
300 to 600	.265	.506	−.260	.116	38.08**
600 to 900	.491	.724	−.293	.146	30.47**

(Compare to Fig. 4).
^ $p = .05$, CI 95%.
* $p < .05$, CI 95%.
** $p < .01$, CI 99%.

showed numerically more fixations on the target in the infelicitous [L + H* no-acc] rather than the felicitous [H* L + H*] condition up to about 400 ms (this early difference did not reach significance, mean difference .034 in the 0–300 ms window, 95%CI = .075)[1]. Instead, the effect of felicitous L + H* in the contrast-on-object context appeared much later, in the 600–900 ms window, where fixation proportions were significantly higher for

felicitous than infelicitous accents. This indicates that felicitous use of L + H* on the noun evoked contrast, and eventually led to more frequent fixations to the target. It is important to note that the shapes of the rising functions are remarkably different between Adjective Contrast conditions and Noun Contrast conditions in Fig. 4. As we mentioned above, the two functions in Adjective Contrast context start rising sharply immediately after the onset of the noun, whereas both functions in Noun Contrast context show a relatively shallow initial slope. The timing of rise in fixations in the Noun Contrast context is comparable to the results in Allopenna et al. (1988) and Dahan et al. (2002), where participants started fixating objects 200–300 ms after the onset of the nouns. (The mean difference scores during the object noun for Adjective Contrast vs. Noun Contrast trials were .174, 99%CI = .100 for the 0–300 ms window, .371, 99%CI = .116 for the 300–600 ms window, and .382, 99%CI = .142 for the 600–900 ms window.) This difference indicates that, regardless of the pitch accent on the adjective, participants began fixating the target cells much earlier when they heard a repeated

---

[1] One anonymous reviewer requested an analysis targeting the region that contains the numerical advantage for the infelicitous accentual pattern. The current analysis windows from 1 to 300 and 300–600 ms divide the region of numerical advantage, such that from 1 to 300 it is paired with a region where no effect is present, and from 300 to 600 ms it is paired with a region that contains a disadvantage for the infelicitous pattern. A targeted analysis showed no significant differences for data in the time window from 160 to 420 ms, the region of the numerical advantage. The back-transformed mean fixation proportions were .187 (infelicitous) and .137 (felicitous), with a mean difference of .05, and a 95% confidence interval of .089 ($F = 1.305$).

noun following an adjective that was different from the immediately preceding ornament than when they heard the same color adjective repeated before a non-repeated noun. This effect is not surprising in the context of a visual search among cells organized by ornament type, because in repeated noun conditions, the location of the target ornament type is easily remembered from the immediately preceding trial. In contrast, repeated color adjective instructions were less informative, as they eliminated the most recently visited ornament cell, but left open ten other possible cells.

The results of Experiment 1 demonstrated that felicitous use of L + H* led to early fixations to the target cells, but only when it marked contrastive status on the color adjective, and not when it was used to express the contrast on the noun. The advantage for targets with felicitous L + H* on the prenominal adjective was observed very early, during the pronunciation of the noun. Fixation proportions in both of the Adjective Contrast conditions increased sharply from noun onset, suggesting that participants were actively searching for the target ornament before they had recognized the object noun—that is, on the basis of information from the adjective. Note that in Dahan et al.'s Experiment 1, fixation proportions began to increase approximately 300 ms after the onset of the target word, even when it was repeated from the previous command (e.g., candle → candle). Based on the assumption about the time requirement for complex visual search tasks (Viviani, 1990), Dahan et al. argue that fixation data begin to reflect the processing of a word about 200–300 ms after its onset. In the present results, however, fixations to the target cells increased from the beginning of the noun even in the Infelicitous Adjective Contrast condition with [H* L + H*] pattern, indicating that listeners returned their gaze to the most recently mentioned ornament cell upon hearing the repeated noun even in the absence of contrastive accent on the preceding color adjective. We interpret the immediate rise in fixations to the target in Adjective Contrast context as the effect of repetition, and tentatively argue that listeners did not need 300 ms to process the repeated word and program their saccades accordingly. Thus, the relatively early rise in fixation proportions in the Felicitous Adjective Contrast condition demonstrates a combined advantage of repetition and the preceding contrastive accent. We attribute the lack of such repetition effect in Dahan et al. (2002) to the overt presence of an unmentioned cohort competitor that drew participants' attention even before they heard the target instruction.

Unlike to the Adjective Contrast conditions, the fixation proportions in the two Noun Contrast conditions did not increase immediately after the onset of the target noun. Instead, fixations to the target cells started to rise at around 300 ms in both felicitous and infelicitous trials when the color was immediately repeated. Based on the assumptions above about the timing of saccade planning and execution for non-repeated objects, we interpret these late increases in fixation as the result of processing the segmental and intonational information in the object nouns. The lack of anticipatory fixations in the Noun Contrast conditions indicates that participants did not plan saccades to the target until they heard the object nouns. We suspect that this difference in the ease of anticipation comes from the way the search task environment was structured. The ornaments were sorted by object types on the grid, and there was no duplication of ornaments (i.e., the same object in the same color) in any cell. Thus repeated noun information provided clear information about the location of the target, while repeated color information did not. In addition, we found no evidence that the combination of repeated color term and *lack* of a contrastive accent presaged an upcoming contrast on the object noun. In sum, the intonational cues of prenominal adjective did not promptly facilitate eye movements when participants needed to wait for the object noun in order to plan a saccade.

We also note that slightly earlier fixations would be predicted for the condition with L + H* on the noun, if felicitous L + H* unconditionally facilitated lexical access. The present data instead suggest that the accentual prominence itself does not trigger earlier outset of word recognition. The timing of fixation rise indicates that felicitous contrastive accent facilitated word recognition once enough segmental information was available. Therefore, participants seem to have selectively tuned to and made use of intonational cues that were informative to performance of the search task in the given environment.

The results of Experiment 1 demonstrated the robust effect of intonational cues that could facilitate anticipatory eye movements. However, a problem remains with the interpretation of data since the felicitous uses of L + H* were always compared against the *infelicitous* use of L + H*. Because contrastive prominence on an immediately repeated word sounds odd in the absence of unusual discourse context, it is possible that the differences shown in Experiment 1 are due to a *delay* in the processing of informational status of the word marked with infelicitous L + H*. We designed Experiment 2 to eliminate this possibility and confirm the facilitative effect of felicitous L + H* during the visual search. In addition, we reasoned that if the presence of an L + H* accent on the adjective was the source of an expectation of contrast, and thus could lead listeners to anticipate the immediate repetition of the noun target, L + H* accents should also function to mislead listeners in a case where the upcoming noun is not a repetition. Experiment 2 was designed to include such misleading trials.

Experiment 2

Experiment 2 was conducted to confirm and explore the primary findings from Experiment 1. First, we addressed whether the early fixations to target objects in the felicitous adjective contrast condition were genuinely due to the anticipatory use of $L + H*$ on the contrasting color adjective. In Experiment 1, we compared felicitous to infelicitous pitch accent sequences for adjective contrast trials (e.g. *blue ball → GREEN ball vs. blue ball → green BALL*), leaving open the possibility that differences were due to processing difficulty caused by the presence of an infelicitous $L + H*$ on the target noun. In the current experiment, we compare felicitous adjective contrast to simple absence of a contrastive accent. In the new *neutral* condition for comparison, the adjective carried a $H*$, and the following noun a downstepped $!H*$ (e.g., *blue ball → GREEN ball vs. blue ball → green ball*), which was the most frequent accent sequence for adjective noun pairs in non-contrastive sequences in our previous production study (Ito et al., 2003). Thus any differences between these conditions should be attributable to the presence or absence of a felicitous $L + H*$. Second, we more rigorously test our contention that $L + H*$ on a color adjective evokes a set of alternatives, increasing listeners' expectations that the most recently mentioned target noun will be repeated in the current utterance. If this is so, not only should a $L + H*$ on a color adjective provide an advantage when the just-mentioned noun is repeated, but it should also *mislead* listeners to fixate on the previously mentioned object cell when the following noun is *not* repeated (as in the sequence, *red angel → BLUE drum*). Thus, when the color adjective carries $L + H*$ in a sequence where the object is *not* repeated, initial fixations should be to the (incorrect) cell of the preceding object (e.g., *angel*), with correct target cell fixated later, based on the noun information. If the contrastive accent on the adjective really guides anticipatory selection of the target noun, we should observe a very early increase of fixations to the just-mentioned object cell triggered by $L + H*$ on the adjective, regardless of the segmental information that identifies the following noun. As a result of accentual

misguidance (a pitch accent-based 'garden path' effect), we predict that fixations to the real target object cell should be distinctly delayed.

Methods

Participants

Thirty-six native speakers of American English were recruited at the Ohio State University. None of them participated in Experiment 1. They received partial credit toward a course requirement for their participation.

Materials

Ornaments were identical to those used in Experiment 1, but different adjective noun combinations were created for use in the conditions of Experiment 2. There were four trees to be decorated, each with 26 ornaments (16 targets and 10 fillers), and two trials in each of the four critical conditions shown in Table 4. Again, each color and each object was mentioned at least once, but no more than three times in each tree. Comparison of the Felicitous and Neutral Adjective Contrast conditions allowed us to test whether the effect of felicitous $L + H*$ is facilitative. In the Felicitous and Infelicitous No Contrast conditions, comparison of early eye movements to the target noun cell with those to the cell mentioned on the immediately preceding trial allowed us to test for 'garden-path' effects of infelicitous $L + H*$. Critical conditions were again balanced for absolute word duration and number of syllables in adjectives and nouns. All words had first-syllable stress. (Mean number of syllables and duration in critically compared conditions: Felicitous: adjective 1.2, 327 ms; noun 1.5, 472 ms; Infelicitous: adjective 1.5, 341 ms; noun 1.6, 483 ms; Adjective contrast: adjective 1.2, 334 ms; noun 1.6, 479 ms; No contrast: adjective 1.5, 335 ms; noun 1.6, 476 ms). As Table 4 shows, we constructed No Contrast conditions so that although neither the adjective nor the noun was immediately repeated from the previous trial, both of these words had been recently used in the instructions sequence, and thus were repeated word mentions. This was done to neutralize any effect

Table 4
Experiment 2: Information status and accent pattern for the eight conditions

Type of mention	Preceding context	Target instruction with accent specification
Critical conditions		
Felicitous Adj Contrast	... green drum →	$BLUE_{L + H*}$ $drum_{no\ accent}$
Neutral Adj Contrast	... green drum →	$blue_{H*}$ $drum_{!H*}$
Felicitous No Contrast	blue candy ... green drum ... red angel →	$blue_{H*}$ $drum_{!H*}$
Infelicitous No Contrast	blue candy ... green drum ... red angel →	$BLUE_{L + H*}$ $drum_{no\ accent}$

Note: Adj, Adjective; N, Noun; → indicates immediate adjacency between the target trial and the preceding trial.

that previous mention of the words, or previous fixation of the relevant ornament cell might have on fixations in these conditions.

The same female speaker recorded auditory stimuli with the identical recording settings. Stimuli were re-recorded until the same two ToBI annotators independently confirmed that they bore the intended intonational patterns. Table 5 shows the mean duration of the verb, the article *the*, the color adjective and the object noun, and the mean peak F_0 values of the adjectives and nouns of the target instructions used in Experiment 2.

The eye-tracking procedure in Experiment 2 was identical to that of Experiment 1. Four grids of ornaments were used to decorate four trees, and participants were given the same instructions for the task and the same practice trials. The only differences

between Experiments 1 and 2 were the order of decoration, the distribution of adjectives and nouns into item pairs in conditions, and the intonational patterns of the instructions. Note, adjective-noun pairs in Experiment 2 were assigned to different conditions than they were in Experiment 1. Thus a comparison examination of like conditions that occur in both experiments (specifically, felicitous and infelicitous $L + H^*$ no accent trials) provides a test of whether the effects in these conditions generalize across different sets of items.

Results and discussion

Due to experimenter error in the sequencing of trials on one of the trees, we had to eliminate one trial in the Infelicitous Adjective Contrast condition from every

Table 5
Experiment 2: Mean duration and F_0 values of the target stimuli

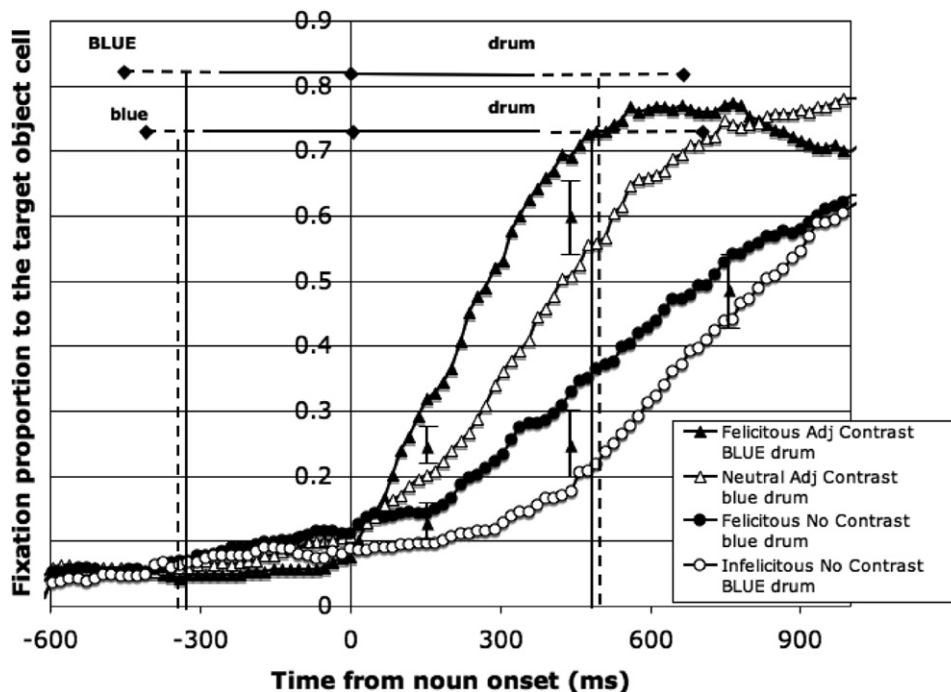| Condition	Verb	*the*	Adj			Noun		
	(ms)	(ms)	(ms)	accent	F_0 (Hz)	(ms)	accent	F_0 (Hz)
Critical conditions								
Felicitous Adj Contrast	265	88	302	$L + H^*$	354	473	—	—
Neutral Adj Contrast	315	85	344	H^*	225	484	!H^*	191
Felicitous No Contrast	325	101	342	H^*	213	493	!H^*	187
Infelicitous No Contrast	334	99	338	$L + H^*$	360	482	—	—

Fig. 5. Experiment 2: Proportion of fixations to the target object cell, aligned from noun onset in the four critical conditions.

participant's data (Neutral Adj Contrast Condition: 36/288 observations missing).

Fig. 5 shows the proportions of fixations to the target cells averaged across participants for the four critical conditions, aligned from the onset of the noun in the same manner as in Fig. 4. Again, these functions show fixation initiation times calculated from saccade onset, and proportions give the number of fixations to target divided by the number of possible fixations during each 17 ms camera cycle (36 Ps × 8 trials per condition = 288, except for Infelicitous Adjective Contrast condition = 252). Statistical analyses consisted of 2 (discourse contrast) × 2 (accent type) repeated measures ANOVAs conducted on arcsine transformed fixation proportions for each 300 ms time window. Table 6 summarizes the results of critical comparisons. The floating bars in Fig. 5 indicate the relevant confidence intervals for the time regions where the mean differences reached statistical significance in the top two comparisons in Table 6. Values of the confidence intervals presented were back-transformed from arcsine-transformed proportions.

Fig. 5 compares the Felicitous and Infelicitous Adjective Contrast conditions, where the same ornament name was immediately repeated to create contrastive status for the color adjective (compare to Fig. 4, Experiment 1). As in Experiment 1, the fixation proportions for both conditions begin to increase from the onset of the target noun, but more sharply in the condition where L + H* felicitously marked the contrastive status of the color adjective than in the neutral condition with H*. The functions diverge very early, within 100 ms of noun onset, and the fixation proportion reaches its peak much earlier in the felicitous condition with L + H*. This pattern of results confirms that the felicitous use of L + H* facilitated visual search by prompting anticipatory identification of the target referent on the basis of information from the adjective. Note that mean word duration for the color adjective was comparable between L + H* and H* in Experiment 2. In fact, the adjectives with L + H* in Felicitous Adjective Contrast condition had numerically shorter average duration than the adjectives with H* in Neutral Adjective Condition (Table 5; vertical reference lines in Fig. 5). Therefore, it is unlikely that the durational differences in the adjectives led the earlier more frequent fixations to the target in the Felicitous Adjective Contrast condition. Correlation analysis confirmed no reliable relation between adjective duration and the first fixation latencies in all four critical conditions ($r = .105$; $.164$; $-.235$; $-.073$ for Felicitous Adjective Contrast, Neutral Adjective Contrast, Felicitous Both Already, and Infelicitous Both Already, respectively.)

Fig. 5 also shows mean fixation proportions for the Felicitous No Contrast condition with [H* !H*] accent pattern and the Infelicitous No Contrast condition with [L + H* no accent] pattern (e.g., *red angel → blue drum*

vs. *red angel → BLUE drum*). Results indicate that an infelicitous contrastive accent did indeed delay fixations to the target noun cell. The fixation proportion was higher for the Felicitous No Contrast condition than for Infelicitous No Contrast condition from the beginning of the noun. While the mean fixation proportion for the felicitous condition rose steadily from about 200 ms after noun onset, that for the infelicitous condition did not begin to rise until approximately 300 ms after noun onset. A consistent gap persisted between the two conditions until beyond fixation proportions of above 50%. As the Felicitous No Contrast condition involved neither a repetition nor any contrastive accent, we may consider its fixation proportion function as the baseline for the visual search task. Thus, the difference between the Felicitous No Contrast condition and the Infelicitous No Contrast condition, which was sustained throughout the noun, indicates delayed processing due to misleading pitch accent information. At the same time, the larger difference between the fixation proportion functions for the Neutral Adjective Contrast and the Felicitous No Contrast conditions indicates the pure effect of repetition.

Figs. 6 and 7 compare fixation proportions within the two No Contrast conditions in order to examine the hypothesis that the presence of a L + H* contrastive accent on the color adjective prompted listeners to actively search for a target object that contrasts in color with the target from the immediately preceding trial. Fixation proportions were averaged across participants for the mentioned target noun and the previous, i.e., incorrect target noun in the infelicitous [L + H* no accent] (Fig. 6) and felicitous [H* !H*] conditions (Fig. 7), both aligned from the onset of the noun. Table 7 summarizes the results of comparisons of mean fixation proportions between the target and the incorrect target for the two conditions. The floating bars in Fig. 6 and 7 indicate the confidence intervals for the regions where the mean differences reached statistical significance shown in Table 7.

Fig. 6 shows fixations to the target ornament cell vs. the immediately preceding ornament cell for instructions with L + H* infelicitously produced on a color adjective that modified a non-repeated object (e.g. *red angel → BLUE drum*). The data show a clear increase in incorrect fixations to the immediately preceding target due to the presence of the infelicitous L + H*. Notice that the increase in incorrect fixations began *before* the onset of the noun and peaked before the noun ended, approximately 250 ms after its onset. This serves as incontestable evidence for the anticipatory effect of contrastive L + H* on the prenominal adjective. The early fixations to the incorrect target (e.g., *angel*) appeared before the object noun could have been recognized, and they continued to increase even after segmental information from the noun might have been used to guide fixations to the

Table 6
Experiment 2: Results comparing fixation proportions ($n = 36$), Subject analysis

Time region (ms)	Mean fixation proportion	Mean fixation proportion	Mean Difference Felicious–Infelicitous	CI	F, $df = 35$
Felicitous Adj Contrast vs. Neutral Adj Contrast					
−600 to −300	.053	.088	−.035	.042	.263
−300 to 0	.058	.091	−.033	.049	1.842
0 to 300	.285	.203	.085	.078	8.774**
300 to 600	.690	.490	.246	.137	24.37**
600 to 900	.799	.760	.063	.114	1.231
Felicitous No Contrast vs. Infelicitous No Contrast					
−600 to −300	.073	.054	.019	.040	.140
−300 to 0	.102	.082	.020	.049	.666
0 to 300	.156	.099	.058	.058	4.073^
300 to 600	.321	.189	.137	.102	7.403*
600 to 900	.522	.434	.100	.114	3.138^
Felicitous Adj Contrast [L + H* no-acc] vs. Infelicitous No Contrast [L + H* no-acc]					
−600 to −300	.053	.054	−.001	.034	.002
−300 to 0	.058	.082	−.024	.048	.985
0 to 300	.285	.099	.189	.078	43.76**
300 to 600	.690	.189	.540	.137	128.4**
600 to 900	.799	.434	.459	.153	71.54**
Felicitous No Contrast [H* !H*] vs. Neutral Adj Contrast [H* !H*]					
−600 to −300	.073	.088	−.015	.038	.008
−300 to 0	.102	.091	.011	.048	.204
0 to 300	.156	.203	−.047	.058	2.674
300 to 600	.321	.490	−.184	.137	13.50**
600 to 900	.522	.760	−.309	.153	31.10**

(Compare to Fig. 5).

^ $p = .05$, CI 95%.
* $p < .05$, CI 95%.
** $p < .01$, CI 99%.

correct target cell. The fact that the increase in fixations to the correct target cell was delayed until after 300 ms into the noun suggests that processing was considerably disrupted by the conflict between the anticipated referent and the incoming noun.

In contrast, Fig. 7 shows no such increase in looks to the immediately preceding target when the same no contrast sequence was produced with a felicitous [H* !H*], dispelling any concern that participants may have used a general strategy of looking back to the most recent cell. The figure instead shows that participants rarely fixated the previous target throughout the felicitously produced target noun phrase.

The results of Experiment 2 confirmed the anticipatory effect of L + H* on the contrastive adjective and demonstrated that this effect was facilitative in nature. That is, listeners planned and executed saccades to the just-mentioned ornament cells earlier when the color adjective carried contrastive prominence L + H* than when it was heard with the less prominent accent H*. In addition, even an infelicitous L + H* on the adjective led to strong anticipation, yielding an increase in incorrect fixations to the immediately pre-

ceding target ornament cell. The timing of the incorrect fixations indicates that listeners planned and executed saccades immediately upon hearing contrastively accented adjectives. It is important to note that these results do not simply confirm the incremental processing of intonational cues shown by Dahan et al. (2002), but also demonstrate how accentual cues interact with the grammatical roles of words to anticipate upcoming input. In both Experiment 1 and 2, fixation patterns showed that listeners quickly integrated the modifier role of the color adjective with its accentual prominence L + H*, assigning contrastive discourse status to the color adjective against the preceding ornament's color, and simultaneously constraining the candidate referent for the upcoming noun to be the same ornament type as the immediately preceding trial. This integration took place rapidly enough to show an increase in fixations to the target from the onset of the noun, indicating that ballistic saccades were planned and executed based solely on the adjectives' grammatical role and their intonational information. The robustness of this anticipatory effect was demonstrated with the infelicitous L + H*
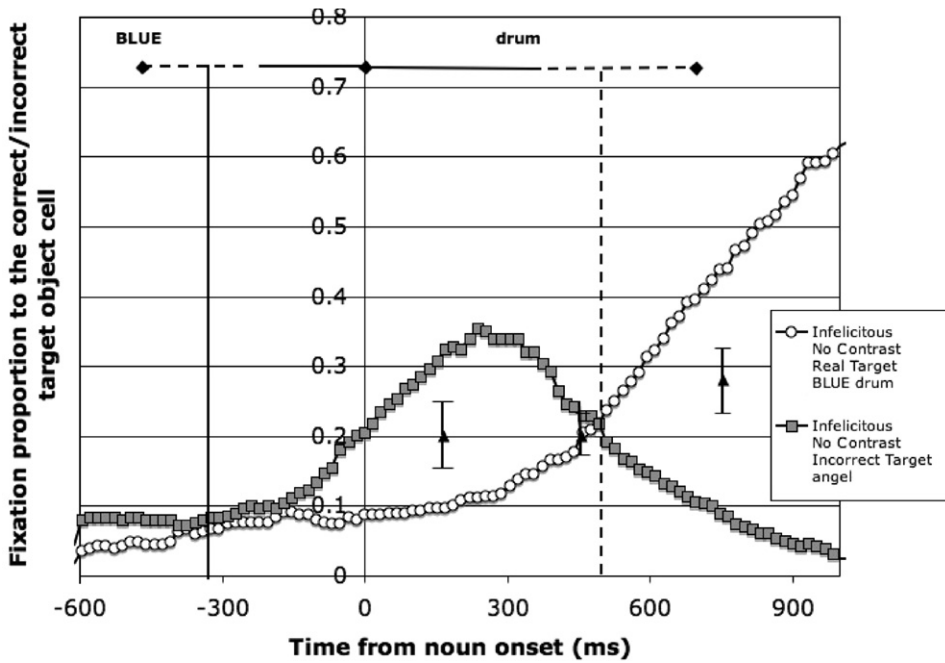
Fig. 6. Experiment 2: Proportion of fixations to the mentioned vs. previous target object cells, aligned from noun onset in the Infelicitous Adjective Contrast [L + H* no accent] condition.
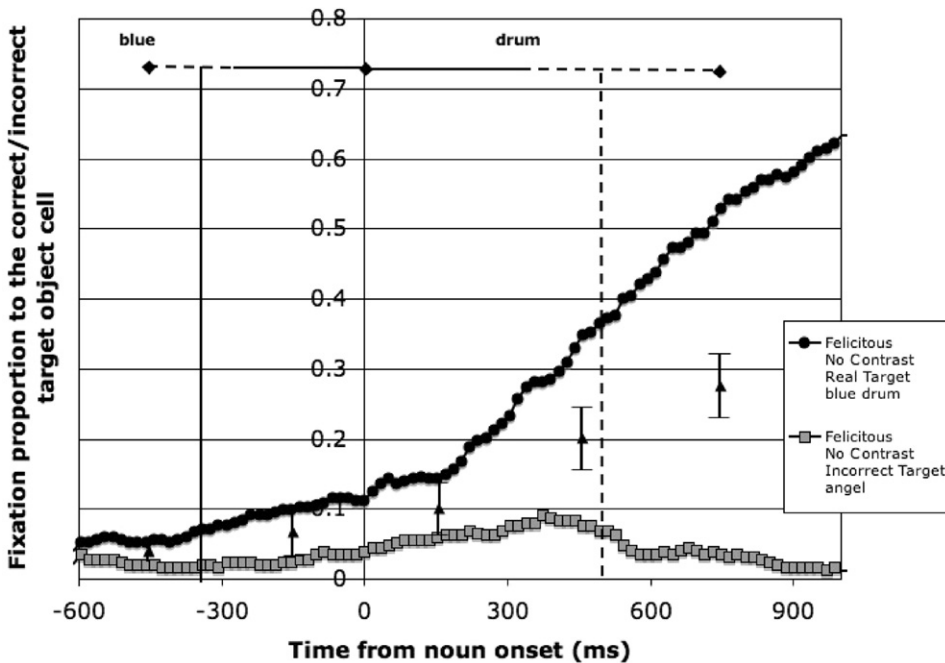

Fig. 7. Experiment 2: Proportion of fixations to the mentioned vs. previous target object cells, aligned from noun onset in the Felicitous No Contrast [H* !H*] condition.

combined with non-repeated objects, which acted like a false alarm to automatically trigger *garden path* eye movements to the incorrect cells.

Although Experiment 1 and 2 confirmed the anticipatory effect of L + H* within a noun phrase, such an effect may have a limit in its scope. Both Dahan et al.

Table 7
Experiment 2: Results comparing fixation proportions ($n = 36$), Subject analysis

Time region (ms)	Mean fixation proportion	Mean fixation proportion	Mean Difference Target–Incorrect target	CI	$F, df = 35$
Infelicitous No Contrast Target vs. Incorrect target					
−600 to −300	.054	.078	−.024	.035	2.035
−300 to 0	.082	.130	−.048	.052	3.410
0 to 300	.099	.290	−.194	.092	32.89[**]
300 to 600	.189	.261	−.074	.068	4.933[*]
600 to 900	.434	.099	.343	.097	95.56[**]
Felicitous No Contrast Target vs. Incorrect target					
−600 to −300	.061	.021	.040	.035	5.257[*]
−300 to 0	.102	.029	.073	.070	8.058[**]
0 to 300	.156	.057	.100	.093	8.768[**]
300 to 600	.321	.074	.250	.091	57.28[**]
600 to 900	.522	.034	.493	.097	206.9[**]

(Compare to Figs. 6 and 7).
 [^]$p = .05$, CI 95%.
 [*] $p < .05$, CI 95%.
 [**] $p < .01$, CI 99%.

and Experiment 1 and 2 in the present study show that the prominent accent $L + H*$ evokes contrast and directs listeners' attention to the alternative(s) available in the given task environment. However, we have not yet addressed the question of the range of $L + H*$ accents in foreshadowing contrast: At what point in the discourse does a prominent $L + H*$ elicit upcoming contrast, thereby leading to anticipatory eye movements? Listeners constantly update their discourse representation to integrate information about the actions and references involved in conversation, and thus a model of intonational processing in discourse must capture the domain of accentual-cue integration. Extending the designs of previous two experiments, Experiment 3 tested the scope of the anticipatory effect of $L + H*$.

Experiment 3

In the tree decoration task used in Experiments 1 and 2, each audio instruction started either with a phrase that specified the direction of decoration such as "*On the right*" and "*To its left*", or with other discourse markers that smoothed the transition between trials such as "*And now*," "*Next*" and "*Finally*." As mentioned earlier, Dahan et al.'s instructions always contained "*Now*" sentence-initially. Such discourse markers (DMs) establish relations between the statements they introduce and the previous utterances, and thus their accentual patterns may affect the process of updating the discourse status of already-mentioned referents and assigning discourse status to upcoming referents.

Using the same tree decoration task, Experiment 3 investigated whether $L + H*$ on the temporal adverbial

DMs *And next, And then*, and *After that* evokes a contrastive interpretation of the upcoming instruction. During a natural discourse or narrative, such DMs signal temporal continuity between the previously described events and those about to be described. Although the speaker may make spontaneous decisions about what to emphasize in the upcoming utterance, the listener may develop an expectation about the focus of upcoming message according to what has been emphasized in the previous utterance and how the DM introduces the next event. Within the task environment of the present study where the action itself is repetitive (i.e., hanging a series of ornaments on the tree), the primary focus of each instruction utterance is the noun phrase that names the target ornament. Thus, the contrastive accent on a DM in a repetitive command "*And THEN, hang the…*" might cue an upcoming contrast between the preceding ornament and the next one. Indeed, we noted an intriguing difference in the timing of effects between Experiments 1 and 2 in Dahan et al. (2002) that may be partly due to a difference in accent patterns on the DM. Although DM accent patterns were not strictly controlled or compared in these studies, the majority of trials in the Experiment 1 deaccented conditions contained a $L + H*$ accent on the word '*Now*,' while in Experiment 2, '*Now*' never carried a $L + H*$. Overall, looks to the preferred object in their Experiment 1 occurred very early, with significant differences at the onset of the target word, while comparable effects in Experiment 2 were not established until approximately 200 ms after the target word onset. Although this difference is anecdotal, it is possible that the DM $L + H*$ might have contributed to the very early differences found in Dahan et al.'s Experiment 1.

In addition, the results of our own Experiment 1 suggest that listeners may be differentially sensitive to accentual cues, depending on their usefulness in a given task environment, such as that for visual search. Comparison of contrast-on-color vs. contrast-on-object trials showed early use of the contrastive L + H* information on color adjectives to guide a search of ornaments sorted by type. However, when the color was repeated and thus the listener had to wait for the noun information to identify the referent, the accentual pattern of the adjective did not have an immediate impact on eye movements. This may indicate that the listeners established very selective use of intonational cues, attending more to those cues most relevant for their particular visual search task. This leaves open the possibility that environment-based strategic tuning to the adjective accent yielded the anticipatory early fixations to the target in Experiment 1 and 2 and the robust 'garden-path' eye movements to the incorrect target due to the contrastive accent in the non-contrastive context in Experiment 2. If the participants were strategically tuning to the color contrast in the present task environment, a contrastive L + H* on a DM may also be interpreted as the signal to a contrast on the color rather than on the object type, and this selective interpretation of the accent may trigger anticipatory fixations to the immediately preceding ornament cell. In other words, the task environment may restrict the domain of contrast signaled by L + H* on a DM, and this contrastive accent at the beginning of utterance may lead to anticipatory eye movement even before the noun phrase information is available. The absence of such anticipatory eye movement would eliminate the possibility that the effect of contrastive accent on the adjective reported in Experiment 1 and 2 were the product of environment-specific attention tuning.

In Experiment 3, DMs had either [L + H* L−H%] or [H* L−H%] intonation on the adverb (e.g., *And NEXT vs. And next*). Table 8 shows the four critical conditions and their accentual patterns. The two accentual patterns were crossed with two accentual patterns on target ornament phrases. Either the ornament was mentioned with a [L + H* no-accent] pattern in a contrast-on-color sequence (e.g., *blue drum → GREEN drum*), or both the adjective and the noun had already been mentioned and appeared again in a non-contrastive sequence, produced with [H* !H*] (e.g., *orange candy → green drum*). Thus, DM accentual patterns were either matched or mismatched with the discourse status and intonational marking of the target ornament. In one of the two matched conditions, both the DM and the color adjective had L + H* ("*First, hang the blue drum*" → "*And THEN, hang the GREEN drum*)*", whereas in the other, neither the DM nor the adjective had contrastive L + H* ("*First, hang the orange candy*" → "*And then, hang the green drum*"). In the two mismatched conditions, only the DM or the adjective had L + H* "(*First, hang the orange candy*" → "*And THEN, hang the green drum*"; "*First, hang the blue drum*"→ "*And then, hang the GREEN drum*"). If L + H* on a DM is sufficient to evoke a contrastive interpretation of the upcoming referent's color, then early fixations to the just-mentioned ornament cell should be observed even before listeners hear the target ornament phrase. On the contrary, if L + H* on a DM is not a strong enough cue to lead to such anticipation, yet plays some role in assigning the contrastive status to the current utterance, its accentual relevance may be evaluated against the informational status of the target ornament NP.

Methods

Participants

Thirty-six native speakers of American English were recruited at the Ohio State University. None of them participated in the previous two experiments. They received partial credit toward a course requirement for their participation.

Design and materials

The same set of ornaments as those in Experiment 1 and 2 were used in Experiment 3, with adjective-noun pairs assigned to conditions in a pattern different from that used in the previous experiments. Each participant decorated four trees, which had 26 ornaments each (16 targets and 10 fillers). Each tree had two trials in each of the four conditions shown in Table 8. Each color and each object was mentioned at least once, but no more than three times in each tree. Conditions were balanced for duration and number of syllables in adjectives and nouns. Duration was balanced for adjective accent

Table 8
Experiment 3: Information status and accent pattern for the four critical conditions

Congruence	L + H* location	Preceding context	DM	Target instruction
Matched	Both	blue drum →	And NEXT$_{L + H*}$	GREEN$_{L + H*}$ drum$_{(no\ accent)}$
Mismatched	Adj	blue drum →	And next$_{H*}$	GREEN$_{L + H*}$ drum$_{(no\ accent)}$
Matched	Neither	orange candy →	And next$_{H*}$	green$_{H*}$ drum$_{!H*}$
Mismatched	DM	orange candy →	And NEXT$_{L + H*}$	green$_{H*}$ drum$_{!H*}$

(Compare to Fig. 8).
Note: Adj, Adjective; DM, Discourse marker.

conditions, but this was not possible for nouns, as unaccented nouns are unavoidably shorter than accented nouns, even in sentence-final positions. All words had first-syllable stress. (Mean number of syllables and duration in accent conditions: L + H* adjectives: 1.2, 324 ms, H* adjectives: 1.3, 329 ms. !H*nouns: 3.3, 519 ms, Unaccented nouns: 3, 465 ms.)

In Experiment 3, the critical comparisons were among the matched and mismatched accentual patterns when the color adjective was contrastive (Matched Both vs. Mismatched Adj), and when there was no immediate contrast (Matched Neither vs. Mismatched DM). Each of the three discourse markers "*And next*," "*And then*," and "*After that*" appeared two to three times in each of the four conditions across the four trees. These discourse markers were produced with either L + H* or H* on the adverb, followed by a L−H% phrasal–boundary tone combination. The discourse markers of the other eight trials on each tree were varied according to the status of each instruction such that they conveyed natural local discourse contexts (e.g., "*At the top*", "*Starting on the right*, "*Moving to the left*", "*Following that*", etc.). The accentual patterns of those DMs were either [H* H−L%] or [L + H* H−L%]. Four of the ten filler instructions had one each of the four accentual patterns of the critical conditions, while the remaining six instructions had [H* H−H%] for their DMs. These accentual patterns were selected based on the filler items in Experiments 1 and 2, so that the variation in the DMs and their accentual patterns were comparable across experiments.

The same female speaker produced the instructions for Experiment 3 with the identical recording setting as in Experiments 1 and 2. The instructions were re-recorded until the same two independent ToBI annotators confirmed the accentual patterns were produced as intended. The intervals between the offset of the DM and the onset of verb were manually edited to fall between 300 and 350 ms (mean 339 ms, stdev 26 ms). The mean duration and the F_0 peak value of the last word of the discourse marker, the color adjective and the object noun are shown in Table 9.

Procedure

The eye-tracking procedure in Experiment 3 was identical to that of Experiments 1 and 2. Four grids of

ornaments were used to decorate four trees, and participants were given the same instructions for the task and the same practice trials. The only differences between Experiment 3 and the preceding two experiments were the distribution of items across conditions, the order of decoration and the intonational patterns of the instructions.

Results and discussion

Results revealed no immediate anticipatory effect of L + H* on the DM itself. Fig. 8 shows the proportions of fixations to the target cells averaged across participants for the four critical conditions, aligned from noun onset. Fixations between the DM and the noun onset are shown with negative time values and vertical lines indicate the mean duration of each word in the instruction as well as the duration of the pause following the DM. The fixation proportion functions before noun onset give no indication of anticipatory eye movements due to contrastive accent on the DM. In all four conditions, participants rarely fixated the target cell before the noun onset. Repeated measures ANOVAs on arcsine transformed proportions showed no effect of DM accent in the six 300 ms windows before the noun onset (the only window of $F > 1$ was −1200 to −150 ms: $F = 1.681$, $p = 2$). The absence of fixations to the target before the noun in the Matched Both trials (e.g., "*And NEXT, hang the GREEN drum*") suggests that a L + H* on a DM did not trigger the anticipation of upcoming color contrast.

In order to confirm this null effect of DM contrastive accent, fixations to the target cell were compared with the fixations to the incorrect target cells (i.e., the preceding trial's target) in the Mismatched DM trials (*orange candy* → "*And NEXT, hang the green drum*"; incorrect target = *candy*) If a L + H* on the DM acted as & *false-alarm* in a way similar to the infelicitous L + H* on the adjective in the non-contrastive trials of Experiment 2, a similar increase in fixations to the incorrect target would be observed before the noun phrase. However, as shown in Fig. 9, the contrastive accent on the DM triggered no garden-path eye movement before the listeners fixated correct target as the noun information became available. In fact, participants rarely looked

Table 9
Experiment 3: Mean duration and F_0 values of the target stimuli

Congruence	L + H* location	Discourse marker			Adjective			Noun		
		(ms)	accent	F0 (Hz)	(ms)	accent	F_0 (Hz)	(ms)	accent	F_0 (Hz)
Matched	Both	627	L + H*	342	304	L + H*	374	511		—
Mismatched	Adj	545	H*	237	344	L + H*	342	420		—
Matched	Neither	489	H*	238	309	H*	219	549	!H*	183
Mismatched	DM	605	L + H*	332	349	H*	236	488	!H*	194

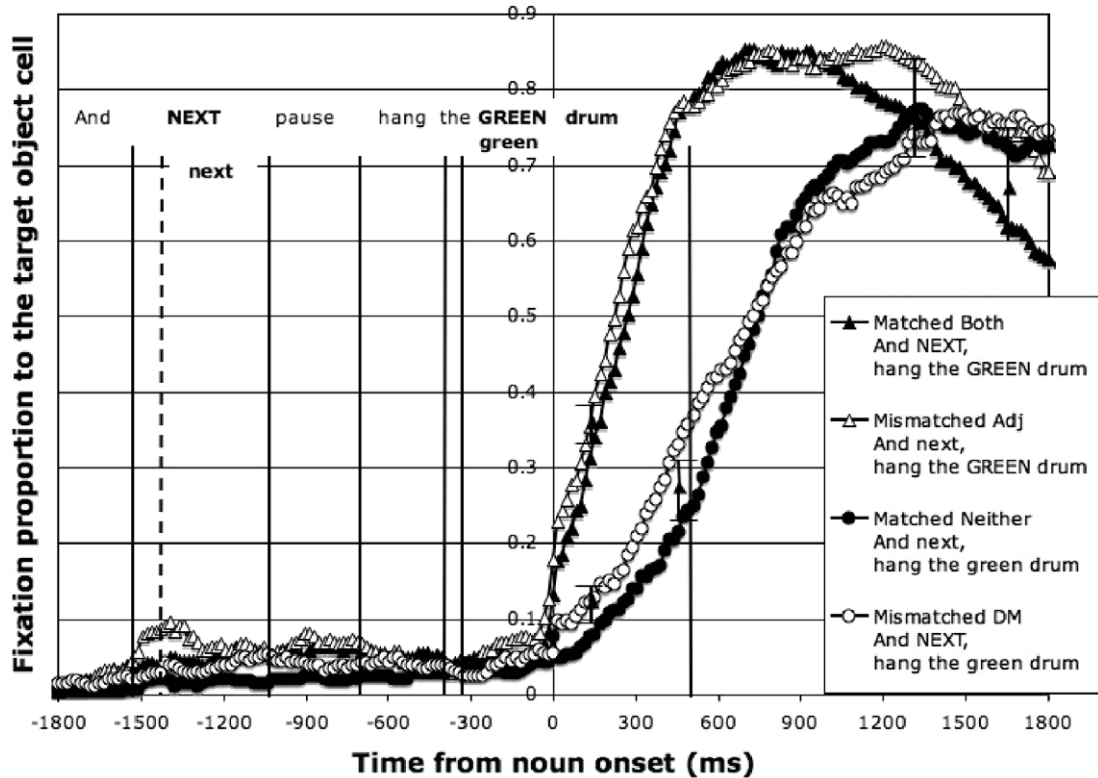*Note:* Adj, Adjective; DM, Discourse marker.

Fig. 8. Experiment 3: Proportion of fixations to the target object cell, aligned from the noun onset in the four critical conditions.

at the previous target throughout these trials. Instead, fixations to the correct target began after the noun increased steadily afterwards. (The results of means comparisons are given in Table 10.) Therefore, it is very unlikely that the participants developed the strategy of using a selective interpretation of L + H* to perform the visual search task in the present experiment.

Although L + H* on the DM did not lead the anticipatory eye movement to the immediately preceding ornament cell, the presence of L + H* may have played some role in foreshadowing the contrast status of upcoming referent. This effect appeared in the much later time region for the comparison between the Matched Both and the Mismatched Adjective trials, and can be seen in Fig. 8. Note that in both conditions where the contrast-on-color was expressed by L + H* on the adjective, the fixation proportions increased sharply right after the onset of the noun, replicating the patterns for Felicitous Adjective Contrast trials in Experiment 1 and 2. This immediate increase in fixation proportions from the beginning of the noun again indicates that the saccades to the target were planned during the preceding adjective carrying a contrastive L + H*. While the fixation proportions continue to rise until about 700–800 ms after the noun onset in both conditions, a remarkable difference

appears as the fixation proportion function declines. As compared to the plateau above .8 line in the Mismatch Adj condition, fixation proportions in the Matched Both condition start dropping sharply after about 1000 ms from the noun onset. This rapid declination in Matched Both trials indicates that the participants' eyes left the target cell quickly after they identified the correct ornament, and this departure was facilitated when the accentual pattern of DM had signaled a contrast than when it had not. Assuming that the decrease in fixations indicates an attention shift from the target cell, we speculate the role of L + H* on the DM as follows. First, contrastive accent on the DM may have prompted a contrast between the preceding and the upcoming referent without restricting the interpretation to the color contrast. Then the L + H* on the color adjective triggered the anticipatory selection of upcoming referent, which initiated the increase in fixations to the previously mentioned target cell. When the contrastive status of the referent was confirmed with the noun information, participants could initiate the action to reach for the ornaments on the board. At the same time, they may have started shifting attention to the tree where they had to place the ornament. We suspect that the presence of early contrastive accent on the DM may
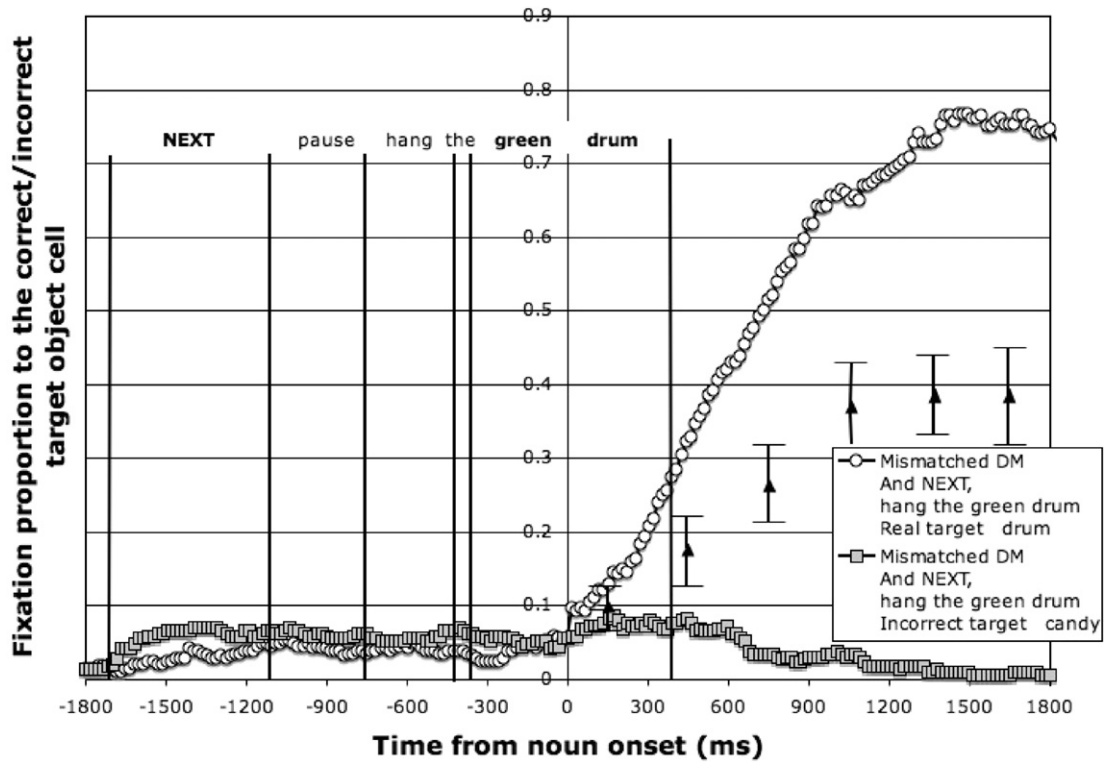
Fig. 9. Experiment 3: Proportion of fixations to the mentioned vs. previous target object cells, aligned from noun onset in the Mismatched DM condition.

Table 10
Experiment 3: Results comparing fixation proportions ($n = 36$), Subject analysis

Time region (ms)	Mean fixation proportion for Mismatched DM		Mean Difference Correct-Incorrect	CI	F, df = 34
	Corrrect target (drum)	vs. Incorrect target (candy)			
orange candy → And NEXT$_{L + H*}$hang the green$_{H*}$ drum$_{!H*}$					
0–300	.129	.073	.056	.044	6.714[*]
300–600	.310	.073	.239	.097	45.81[**]
600–900	.523	.041	.487	.104	178.2[**]
900–1200	.671	.030	.649	.119	263.2[**]
1200–1500	.754	.014	.745	.100	528.1[**]
1500–1800	.806	.008	.801	.131	373.9[**]

(Compare to Fig. 9).
  [^]$p = .05.$, CI 95%.
  [*] $p < .05$, CI 95%.
  [**] $p < .01$, CI 99%.

have facilitated the confirmation of the contrastive status of the referent, and thus speeded the attention shift.

Posthoc means comparisons were conducted between critical conditions in the time windows beginning at -300 ms. Mean fixation proportions in the Mismatched Adj condition were significantly higher than those for the Matched Both conditions in the 0–300 ms window as shown in Table 11. This result was unexpected, but it is unlikely that the narrow gap between the initial rise in these two functions indicates a genuine difference in processing ease due to the accentual property of DM. Instead, we view the sharp rises in the two functions together as the demonstration of a ceiling on the anticipatory effect of contrastive accent on the adjective.

Table 11
Experiment 3: Results comparing fixation proportions ($n = 36$), Subject analysis

Time region (ms)	Mean fixation proportion	Mean fixation proportion	Sine Mean Difference Matched-Mismatched	CI	F, df = 34
Matched L + H* on Both vs. Mismatched L + H* on Adj					
−300 to 0	.099	.074	.025	.029	1.358
0 to 300	.324	.382	−.062	.048	6.545*
300 to 600	.722	.757	−.053	.085	1.629
600 to 900	.874	.879	−.009	.092	.027
900 to 1200	.865	.874	−.018	.123	.091
1200 to 1500	.799	.871	−.132	.129	4.297*
1500 to 1800	.678	.793	−.170	.115	9.149**
Matched L + H* on Neither vs. Mismatched L + H* on DM					
−300 to 0	.116	.044	.072	.039	4.271**
0 to 300	.080	.129	−.049	.049	4.042*
300 to 600	.207	.310	−.106	.085	6.475*
600 to 900	.491	.523	−.037	.110	.457
900 to 1200	.709	.671	.052	.122	.736
1200 to 1500	.771	.754	.026	.128	.166
1500 to 1800	.760	.806	−.074	.115	1.720

^ $p = .05$., CI 95%.
* $p < .05$, CI 95%.
** $p < .01$, CI 99%.

Interestingly, we found an early and continuing series of significant advantages for the Mismatched DM condition over the Matched Neither condition in −300–0, 0–300, and 300–600 ms windows (see Table 11). The Mismatched DM condition (e.g., "*Hang the orange candy. And NEXT, hang the green drum*"), is the condition that is intonationally most like Dahan et al.'s (2002) Experiment 1 deaccented condition, where the sentence pair was, "*Put the candle above the star. NOW put the can ...*" However, the two conditions differ substantially in the informational content of their accent patterns. In Dahan et al.'s materials, the felicitous use of the contrastively accented DM followed by a deaccented target noun involved the repetition of the immediately preceding noun target. In the current experiment, neither the adjective nor the noun was repeated. We speculate that the presence of L + H* on the DM, when compared to the absence of such contrastive cue, may have simply had a generalized attentional effect, speeding processing of the repeated target in Dahan et al. (2002) as well as the non-contrastive target in the current experiment, resulting in an effect that surfaced once the noun phrase information became available. This effect may be present but not measurable in our Matched Both condition due to the ceiling effect discussed above. Although this is a plausible account of the present outcome, further research is required to explore the general function of the contrastive accent on the discourse marker.

The results of Experiment 3 demonstrated that contrastive accent on a DM does not evoke the strategic expectation of a contrast in the upcoming utterance. Participants in the present study did not develop the use of a strategic interpretation of contrastive accent even with the repetitive commands to search ornaments sorted by object types, which may have highlighted the possible advantageous use of L + H* for the adjective but not for the noun. Although L + H* on the DM may have signaled upcoming contrast, or may have simply provided an additional attention-orienting cue, it did not specify the type of contrast, and thus it did not trigger early fixations to the immediately preceding ornament cell. The present results suggest that a contrastive accent on the DM is evaluated when the informational status of the referent becomes available with the noun phrase naming the referent. The early prompt of the contrast seems to facilitate the attentional shift after the contrast is confirmed at the later point in the utterance.

General discussion

We have presented three eye-tracking experiments examining the role of pitch accent in discourse comprehension during a relatively complex real-world visual search task. Participants followed pre-recorded instructions for a tree decoration task that necessitated a search through over 40 to 52 small ornaments in an organized

grid display. Each instruction was followed by a visual search and the hanging of the selected ornament (for some, a relatively delicate manual process). This naturalistic task captured participants' attention, and combined with head mounted eye movement monitoring, allowed us to measure implicitly the time course of listeners' use of intonational patterns to anticipate and recognize discourse referents. Even with the inclusion of intonationally infelicitous trials, post-experiment debriefing interactions with participants informed us that they uniformly did not consider accents on words as a possible focus of the study, with few even attending to intonation. Many participants thought that the experiment was testing their responses to color distribution or their memory of the objects' locations (ornament boards vs. trees).

Experiment 1 compared felicitous and infelicitous use of L + H*. When L + H* felicitously marked contrast on a color adjective modifying a repeated noun ("*First hang the green drum.*" → "*Next, hang the BLUE drum.*"), fixation proportions to target cells increased more quickly than when L + H* infelicitously marked the immediately repeated noun ("*Hang the blue onion*" → "*Next, hang the blue DRUM*"). In contrast, felicitous L + H* on the noun ("*Hang the blue onion*" → "*Then, hang the blue DRUM*") did not lead to an early increase in fixations to the target as compared to infelicitous L + H* on the immediately repeated color adjective ("*First, hang the blue onion*" → "*Then, hang the BLUE drum*"). The effect of felicitous L + H* on the noun appeared after the noun's segmental information was fully available. The results of Experiment 1 suggested that listeners may have 'tuned' to tonal cues that were relevant to the task environment, where no two objects within a cell had the same color and different objects of the same color were distributed across cells. Within this visual context, a contrastive accent on the non-repeated prenominal adjective was useful to predict the upcoming object type, and participants were able to rapidly integrate pitch accent information to guide visual search. In contrast, the pitch accent pattern on repeated adjectives did not lead to an immediate difference in eye movements. While these results demonstrated clear effects of pitch accent type and location on establishing a contrast set from which to choose the upcoming referent during discourse comprehension, they could not establish whether the effects were due to a processing advantage conferred by felicitous use of L + H*, or instead to disruption from infelicitous use.

Experiment 2 confirmed that the presence of L + H* (as opposed more neutral H*) on a non-repeated adjective provided a processing advantage, triggering the selection of a candidate for the following noun. The anticipatory nature of the effect was demonstrated by 'garden-path' eye movements when the pitch accent pattern on a non-repeated adjective was misleading. Upon hearing a contrastive L + H* accent on the color *(red angel → GREEN drum)*, participants immediately fixated the incorrect, previously-mentioned object cell. The effect was clearly due to the accented adjective, because it began before segmental information for the noun became available. Importantly, the proportion of fixations to the incorrect object cell continued to increase from before noun onset until 300 ms into the noun. That is, listeners' eyes continued to be drawn to the anticipated object even as they listened to conflicting segmental information. Such incorrect initial fixations were not observed in the absence of L + H* on the adjective modifying the non-contrastive referent, ruling out the possibility that participants had simply developed a strategy of looking back to the previous cell.

In Experiment 3, L + H* was placed on discourse markers (e.g., *And NEXT*) that preceded the repetitive command "*. . . hang the . . .*" to test whether participants had developed a selective interpretation of the presence of a L + H* as 'contrast-on-color' as a strategy to perform the visual search task. No sign of specific anticipation due to the contrastive accent on a DM was observed in the eye fixation patterns before the target noun phrase. Instead, when a L + H* on the DM prompted an upcoming contrast on the color, the attention shift (presumably to the tree) seemed to be speeded. In addition, L + H* on the DM may have generally drawn attention to the target noun phrase, as indicated by early more frequent fixations to the non-contrastive target.

Our results are consistent with those from previous eye movement studies that have demonstrated very early use of prosodic information during real-time processing, with anticipatory looks to targets made on the basis of intonational cues and before confirming lexical information (e.g., Snedecker & Trueswell, 2003; Dahan et al., 2002). The effects support a model of spoken language processing that assumes immediate, parallel processing of segmental and suprasegmental information such as pitch accent, despite the latter's distribution across multiple phonemes in the speech stream. Our participants showed immediate sensitivity to the presence of and type of pitch accent, integrating it with information from the discourse representation, and using it to speed ongoing identification of the object of visual search. Taken together, the results of the three experiments presented here also make two novel contributions to our understanding of how listeners use pitch accent information to establish referential domains during discourse comprehension. Specifically, the work demonstrates how the immediate integration of pitch accent information into the discourse representation

can generate an expected referent, and also how grammatical roles of words and referential context constrain the domain of accent-based referential resolution.

First, a L + H* on a prenominal adjective immediately evoked a contrast between the accented discourse entity (i.e., the accented color) and the most salient entity that shared the same grammatical role in the discourse background (i.e., just mentioned color). Simultaneously, this contrastive link between the two prenominal modifiers evoked a mapping between the two modified nouns, projecting a specific candidate in the discourse foreground. The data indicate that eye movements to incorrect targets were planned based on the accentual information of the prenominal modifier, and executed even in the presence of conflicting segmental information. Because saccades are ballistic motor movements, they cannot be reprogrammed after they are initiated. Thus, they may not reflect the exact time course of processing of conflicting speech signals. However, the finding that the misleading intonational cue delayed fixations to the object even while its phonemic information was available strongly suggests that pragmatic information from contrastive accent is processed immediately, on par with other acoustic information in the speech stream. Not only did pitch accent produce an incremental update of the informational status of the currently processed word, but it also initiated predictive lexical access.

Second, our results show that the effect of accentual cues seems to be constrained by the discourse/ grammatical role of the word conveying the accent. Although we found a robust anticipatory effect with the contrastive accent on the adjective, the same prominent accent did not produce an equivalent effect in the discourse marker position. Although speakers often use contrastive accent on DMs to draw the listener's attention to a specific (or maybe an entire) part of upcoming message, because DMs are largely independent of the syntactic structure of the following utterance, listeners have insufficient information to generate specific hypotheses about upcoming referent possibilities. In contrast, a determiner-adjective sequence provides enough information for the listener to project a head noun. Presumably, the accentual cues can be integrated faster when they accompany words whose grammatical roles constrain the upcoming informational structure. At the same time, the accentual cues associated with particular grammatical roles may be constrained by referential context, as demonstrated by the weaker effect of accentual cues on the repeated adjective in the present study's search task environment. Thus, our present results suggest that both referential context and grammatical structure may define the domain and the strength of accen-

tual effects. Further research is needed to explore both the scope of referential constraint and the scope of syntactic constraint on the effect of accentual cues during speech comprehension.

The present study demonstrated robust, pervasive effects of contrastive accent on the processing of discourse referents during visual search. The current experiments stand in contrast to previous work by Sedivy et al. (1999), which failed to demonstrate any effect of intonationally marked contrast with substantially similar materials. In their Experiment IB, Sedivy et al. (1999) monitored participants' eye movements while they followed auditory commands to touch one of four objects: a minimal contrast pair with a pronominal modifier (e.g., a pink comb and a yellow comb), a competitor that shared the contrast property (e.g., a yellow bowl), and a distracter that did not share the contrast (e.g., a metal knife). For each critical trial, participants first heard an instruction that mentioned one of the minimal-pair objects (e.g., "*Touch the pink comb*"). The following instruction mentioned either the counterpart of the minimal pair (e.g., the yellow comb) or the competitor (e.g., the yellow bowl), and modifiers were produced either with L + H* or H*. Results showed that modifiers were immediately interpreted as contrastive, with fast eye movements to the minimal-pair counterparts. However, no effect of the H* vs. L + H* accentual difference was observed.

We notice three major differences between Sedivy et al.'s work and that presented here, which may have led to the difference in findings across the two studies: display complexity, informativeness of the adjective modifier in the discourse context, and consistency of phonological information in the spoken materials. First, display complexity differed substantially across the studies, with four objects in Sedivy et al., and more than ten times that number in the current experiments. In a set of four, the relationship between the minimal pair and competitor objects would be salient to participants, who observed each display change for about 20 s before the initial instruction on each trial. Thus, the preview and relatively simple display may have allowed listeners to establish a double contrast (pink/yellow comb and yellow comb/bowl). We suspect that this display-oriented overt contrast, rather than the general 'contrastive interpretation' of the adjectival modifier (Sedivy et al., 1999, pl27) led to a ceiling effect (mean fixation latencies ranged from 270 to 281 ms). In the present experiments, ornaments were sorted into 11 cells, and each cell contained three to five ornaments. This visual complexity engaged participants in visual search rather than simple object selection from a known field. In addition, in the current study there was no display-oriented referential bias, as there were multiple competitors of the same

color across cells. In other words, participants had no way of guessing the next target ornament based on non-linguistic contrast within the display. In order to make a direct comparison of the timing of effects across the two studies, we calculated the average latency of first fixations for each condition in Experiment 1 and 2. The average first fixation latencies for the felicitous L + H* conditions in Experiment 1 and 2 were 407 and 388 ms, respectively. We argue that the approximate 120 ms difference in the fixation latency across the two studies was driven by the inequality in display complexity and in non-linguistic contextual bias. The relatively slow increases in fixations in the present experiments may be due also to the visual complexity in the experimental setup. (For the detailed discussion on the relation between the visual complexity, preview time and fixations in scene perception, see Henderson & Ferreira, 2004).

A second difference between studies is in the informativeness of the prenominal modifiers used. Sedivy et al. (1999) posit that their adjectives were interpreted contrastively regardless of their accents because the wide range of modifiers used increased their informativeness (objects were described with size, color and material terms, and also presented unmodified). They argue further that the presence of a modifier in the initial instruction (pink comb) may have drawn extra attention to the contrasting modified object (yellow comb) during the initial instruction. Thus excessive informativeness of modifiers across the experiment led to very fast eye movements to the minimal pair object (yellow comb) as compared to the competitor (yellow bowl)—a ceiling effect that obscured the effect of felicitous contrastive accent. We instead attribute the ceiling effect to the varied informativeness of modifier within a display - while the modifier was informative for minimal pair members (pink/yellow comb), the modifier was also used in a less informative manner to describe the single object that shared color with a member of the minimal pair (yellow bowl). When the modifier conveyed unnecessary or confusing information about the target, participants may have interpreted it as the modifier of the contrastive object (e.g., the yellow comb) which required the modification in order to be distinguished from the other member of the pair. We suggest that the overt visual contrast led to frequent looks to the contrastive object regardless of the accentual property on the modifier. In the present study, the informativeness of color modifier was consistent. There were multiple ornaments of multiple colors, and no ornament could be singled out without a color modifier. Therefore, the color modifier in each instruction was equally informative, allowing the examination of the effect of accent uninfluenced by differences in informativeness.

Finally, the phonetic consistency of instructions is crucial for examining prosodic effects. Sedivy et al. (1999) gave the oral instructions by reading aloud the script, pronouncing L + H* accents in the 'Stress' condition and H* in the 'No stress' condition. Since no prosodic transcriptions or acoustic analysis of the instructions were provided, it is difficult to ascertain the phonetic distinction between the two conditions. Pitch range and speech rate are highly variable within speakers, and it is not easy to consistently pronounce equivalent tunes across utterances. It is possible that some L + H* instructions of Sedivy et al. (1999) were produced with less (or more) prominent accents than our speech materials. The present study employed careful ToBI analysis for screening the speech stimuli. Although the debate remains open over phonological distinction between L + H* and H* (Ladd & Morton, 1997; Ladd & Schepman, 2003), we ensured that the accentual values of our speech stimuli were phonetically distinct across conditions. Although we do not wish to devalue the importance of investigating the online processing of non-cardinal accents produced in spontaneous speech, we feel it is critical to provide phonetic and phonological analysis of experimental materials used in research on processing and intonation. Current work in our laboratory examines the production and detection of categorical boundaries across different accent types in online spontaneous dialogue comprehension.

The research presented here adds to a growing body of work that employs natural tasks to structure the attention and intentions of interlocutors during discourse production and comprehension. Here, we have successfully used a visual search task and head mounted eyetracking to examine the time-course of contrastive pitch accent use. We feel that the combination of such measures and tasks with careful phonetic and phonological analyses of spoken materials will lead to an accurate characterization of the use of intonation in language processing.

Appendix A. Experimental trails for Experiment 1

Trial	Condition	Adjective	Noun	Accent Pattern
Tree 1				
1	Filler	white	tree	H* !H*
2	Both Initial	red	drum	H* !H*
3	Both Initial	grey	ball	H* !H*
4	Filler	purple	tree	H* !H*
5	Filler	white	snowman	H* !H*
6	Initial Adj	blue	ball	H* !H*
7	Initial N	grey	candy	H* H*
8	Infel Adj Cont	gold	CANDY	H* L + H*
9	Filler	white	star	H* H*
10	Initial N	red	stocking	H* H*
11	Filler	yellow	tree	H* !H*
12	Both Already	blue	drum	H* !H*
13	Infel N Cont	BLUE	egg	L + H* no-acc
14	Fel Adj Cont	SILVER	egg	L + H* no-acc
15	Filler	purple	star	H* !H*
16	Initial Adj	brown	ball	H* !H*
17	Fel N Cont	brown	ANGEL	H* L + H*
18	Infel Adj Cont	green	ANGEL	H* L + H*
19	Filler	purple	snowman	H* !H*
20	Both Already	green	egg	H* !H*
21	Fel N Cont	green	ONION	H* L + H*
22	Fel Adj Cont	ORANGE	onion	L + H* no-acc
23	Infel N Cont	ORANGE	bell	L + H* no-acc
24	Filler	yellow	star	H* !H*
Tree 2				
1	Filler	purple	tree	H* !H*
2	Both Initial	silver	ball	H* !H*
3	Both Initial	grey	angel	H* !H*
4	Filler	white	light bulb	H* !H*
5	Filler	yellow	snowman	H* !H*
6	Initial Adj	gold	angel	H* !H*
7	Filler	white	tree	H* !H*
8	Initial N	grey	stocking	H* H*
9	Infel Adj Cont	brown	STOCKING	H* L + H*
10	Filler	purple	light bulb	H* !H*
11	Initial N	silver	egg	H* H*
12	Both Already	gold	ball	H* !H*
13	Fel N Cont	gold	CANDY	H* L + H*
14	Infel Adj Cont	green	CANDY	H* L + H*
15	Filler	purple	snowman	H* !H*
16	Initial Adj	blue	angel	H* !H*
17	Infel N Cont	BLUE	bell	L + H* no-acc
18	Fel Adj Cont	ORANGE	bell	L + H* no-acc
19	Filler	yellow	light bulb	H* !H*
20	Both Already	orange	candy	H* !H*
21	Infel N Cont	ORANGE	onion	L + H* no-acc
22	Fel Adj Cont	RED	onion	L + H* no-acc
23	Fel N Cont	red	DRUM	L + H* no-acc
24	Filler	white	snowman	H* !H*
Tree 3				
1	Filler	yellow	star	H* !H*
2	Both Initial	brown	bell	H* !H*
3	Both Initial	orange	stocking	H* !H*
4	Filler	purple	light bulb	H* !H*
5	Initial Adj	silver	stocking	H* !H*
6	Initial N	orange	candy	H* H*

Appendix A (*continued*)

Trial	Condition	Adjective	Noun	Accent Pattern
7	Infel Adj Cont	blue	CANDY	H* L + H*
8	Filler	white	snowman	H* !H*
9	Filler	yellow	light bulb	H* !H*
10	Initial N	brown	angel	H* H*
11	Both Already	silver	bell	H* !H*
12	Infel N Cont	SILVER	onion	L + H* no-acc
13	Fel Adj Cont	GREY	onion	L + H* no-acc
14	Filler	purple	snowman	H* !H*
15	Filler	white	light bulb	H* !H*
16	Initial Adj	green	stocking	H* !H*
17	Fel N Cont	green	DRUM	H* L + H*
18	Infel Adj Cont	red	DRUM	H* L + H*
19	Filler	yellow	snowman	H* !H*
20	Filler	white	star	H* !H*
21	Both Already	red	onion	H* !H*
22	Fel N Cont	red	EGG	H* L + H*
23	Fel Adj Cont	GOLD	egg	L + H* no-acc
24	Infel N Cont	GOLD	ball	L + H* no-acc
Tree 4				
1	Filler	white	light bulb	H* !H*
2	Both Initial	blue	egg	H* !H*
3	Both Initial	green	bell	H* !H*
4	Filler	yellow	star	H* !H*
5	Filler	purple	tree	H* !H*
6	Filler	WHITE	tree	L + H* no-acc
7	Initial Adj	brown	bell	H* !H*
8	Initial N	green	angel	H* H*
9	Infel Adj Cont	orange	ANGEL	H* L + H*
10	Initial N	blue	onion	H* H*
11	Filler	yellow	light bulb	H* !H*
12	Both Already	brown	egg	H* !H*
13	Fel N Cont	brown	DRUM	H* L + H*
14	Infel Adj Cont	red	DRUM	H* L + H*
15	Filler	white	star	H* !H*
16	Initial Adj	gold	bell	H* !H*
17	Infel N Cont	GOLD	ball	L + H* no-acc
18	Fel Adj Cont	GREY	ball	L + H* no-acc
19	Filler	yellow	tree	H* !H*
20	Filler	purple	light bulb	H* !H*
21	Both Already	grey	drum	H* !H*
22	Infel N Cont	GREY	stocking	L + H* no-acc
23	Fel Adj Cont	SILVER	stocking	L + H* no-acc
24	Fel N Cont	silver	CANDY	H* L + H*

Appendix B. Experimental trials for Experiment 3

Trial	Condition	DM and its accent pattern	Target Adj + N and its accent pattern	
Tree 1				
1	Filler	At the top$_{H* L–H\%}$	purple$_{H*}$	star$_{!H*}$
2	Filler	On the LEFT$_{L + H* L–H\%}$	white$_{H*}$	tree$_{!H*}$
3	Both Initial	In the middle$_{H* H–L\%}$	green$_{H*}$	bell$_{!H*}$
4	Initial N Cont	On the RIGHT$_{L + H* H–L\%}$	green$_{H*}$	EGG$_{L + H*}$

Appendix B (*continued*)

Trial	Condition	DM and its accent pattern	Target Adj + N and its accent pattern	
5	Filler	First$_{H*\ H-H\%}$	yellow$_{H*}$	drum$_{!H*}$
6	Filler	Moving to the right$_{H*\ H-H\%}$	purple$_{H*}$	SNOWMAN$_{L+H*}$
7	Both Initial	Following THAT$_{L+H*\ H-L\%}$	gold$_{H*}$	stocking$_{!H*}$
8	Filler	And then$_{H*\ H-H\%}$	white$_{H*}$	star$_{!H*}$
9	Initial Adj	To the right of THAT$_{L+H*\ H-L\%}$	silver$_{H*}$	bell$_{!H*}$
10	Mismatch Adj	And then$_{H*\ L-H\%}$	ORANGE$_{L+H*}$	bell$_{noacc}$
11	Filler	FIRST$_{L+H*\ H-L\%}$	gold$_{H*}$	candy$_{!H*}$
12	Filler	Moving to the left$_{H*\ H-H\%}$	yellow$_{H*}$	tree$_{!H*}$
13	Initial N	Following THAT$_{L+H*\ H-L\%}$	green$_{H*}$	angel$_{H*}$
14	Match Both	And NEXT$_{L+H*\ L-H\%}$	BROWN$_{L+H*}$	angel$_{noacc}$
15	Initial N Cont	To the left of that$_{H*\ H-L\%}$	brown$_{H*}$	BALL$_{L+H*}$
16	Mismatch Adj	After that$_{H*\ L-H\%}$	GREY$_{L+H*}$	ball$_{noacc}$
17	Filler	Finally$_{H*\ H-H\%}$	purple$_{H*}$	TREE$_{L+H*}$
18	Filler	The first ornament is$_{H*\ H-H\%}$	yellow$_{H*}$	star$_{!H*}$
19	Match Neither	And next$_{H*\ L-H\%}$	green$_{H*}$	stocking$_{!H*}$
20	Match Both	After THAT$_{L+H*\ L-H\%}$	RED$_{L+H*}$	stocking$_{noacc}$
21	Initial Adj	Moving to the left$_{H*\ H-L\%}$	blue$_{H*}$	egg$_{!H*}$
22	Mismatch DM	And THEN$_{L+H*\ L-H\%}$	silver$_{H*}$	candy$_{!H*}$
23	Initial N	To the left of that$_{H*\ H-L\%}$	orange$_{H*}$	onion$_{H*}$
24	Mismatch DM	And NEXT$_{L+H*\ L-H\%}$	red$_{H*}$	drum$_{!H*}$
25	Match Neither	After that$_{H*\ L-H\%}$	blue$_{H*}$	ball$_{!H*}$
26	Filler	The last ornament is$_{H*\ H-L\%}$	white$_{H*}$	snowman 1$_{H*}$
Tree 2				
1	Filler	At the very top$_{H*\ H-H\%}$	white$_{H*}$	tree$_{!H*}$
2	Both Initial	On the RIGHT$_{L+H*\ L-H\%}$	silver$_{H*}$	egg$_{!H*}$
3	Mismatch Adj	And then$_{H*\ L-H\%}$	GREEN$_{L+H*}$	egg$_{noacc}$
4	Both Initial	To the left of that$_{H*\ H-L\%}$	red$_{H*}$	drum$_{H*}$
5	Filler	So FIRST$_{L+H*H-L\%}$	purple$_{H*}$	light bulb$_{!H*}$
6	Filler	To the right of that$_{H*\ H-H\%}$	yellow$_{H*}$	bell$_{H*}$
7	Initial N	Following that$_{H*\ H-L\%}$	red$_{H*}$	onion$_{H*}$
8	Match Neither	And next$_{H*\ H-H\%}$	green$_{H*}$	drum$_{!H*}$
9	Initial N Cont	Moving to the RIGHT$_{L+H*\ H-L\%}$	green$_{H*}$	CANDY$_{L+H*}$
10	Match Both	After THAT$_{L+H*\ L-H\%}$	GOLD$_{L+H*}$	candy$_{noacc}$
11	Filler	FIRST$_{L+H*\ H-L\%}$	purple$_{H*}$	BALL$_{L+H*}$
12	Initial Adj	To the right of that$_{H*\ H-L\%}$	orange$_{H*}$	bell$_{!H*}$
13	Filler	Following that$_{H*\ H-H\%}$	yellow$_{H*}$	light bulb$_{H*}$
14	Filler	Moving to the right$_{H*\ H-L\%}$	white$_{H*}$	snowman$_{H*}$
15	Mismatch DM	And THEN$_{L+H*\ L-H\%}$	orange$_{H*}$	onion$_{H*}$
16	Match Both	And NEXT$_{L+H*\ L-H\%}$	BLUE$_{L+H*}$	onion$_{noacc}$
17	Filler	Finally in this row$_{H*\ H-H\%}$	purple$_{H*}$	tree$_{!H*}$
18	Filler	FIRST$_{L+H*\ H-L\%}$	brown$_{H*}$	drum$_{!H*}$
19	Initial N Cont	Moving to the left$_{H*\ H-L\%}$	brown$_{H*}$	STOCKING$_{L+H*}$
20	Filler	After that$_{H*\ L-H\%}$	RED$_{L+H*}$	stocking$_{noacc}$
21	Initial N	To the left of THAT$_{L+H*\ H-L\%}$	blue$_{H*}$	angel$_{!H*}$
22	Mismatch Adj	And next$_{H*\ L-H\%}$	GREY$_{L+H*}$	angel$_{noacc}$
23	Filler	To the left of that$_{H*\ H-H\%}$	yellow$_{H*}$	SNOWMAN$_{L+H*}$
24	Mismatch DM	After THAT$_{L+H*\ L-H\%}$	silver$_{H*}$	ball$_{!H*}$
25	Match Neither	After then$_{H*\ L-H\%}$	gold$_{H*}$	angel$_{!H*}$
26	Filler	Finally$_{H*\ H-H\%}$	purple$_{H*}$	snowman$_{!H*}$
Tree 3				
1	Filler	At the top$_{H*\ L-H\%}$	purple$_{H*}$	snowman$_{!H*}$
2	Both Initial	Starting on the RIGHT$_{L+H*\ L-H\%}$	silver$_{H*}$	bell$_{!H*}$
3	Filler	In the middle$_{H*\ L-H\%}$	white$_{H*}$	light bulb$_{!H*}$
4	Filler	On the left$_{H*\ H-H\%}$	YELLOW$_{L+H*}$	light bulb$_{noacc}$
5	Initial N	Starting on the left$_{H*\ H-L\%}$	silver$_{H*}$	onion$_{H*}$
6	Mismatch Adj	And then$_{H*\ L-H\%}$	BLUE$_{L+H*}$	onion$_{noacc}$
7	Filler	And NEXT$_{L+\ H*L-H\%}$	white$_{H*}$	star$_{!H*}$
8	Both Initial	To the right of that$_{H*\ H-L\%}$	brown$_{H*}$	drum$_{!H*}$

Appendix B (*continued*)

Trial	Condition	DM and its accent pattern	Target Adj + N and its accent pattern	
9	Match Both	After THAT$_{L + H*\ L-H\%}$	RED$_{L + H*}$	drum$_{noacc}$
10	Inittial N Cont	Following THAT$_{L + H*\ H-L\%}$	red$_{H*}$	BALL$_{L + H*}$
11	Filler	Starting on the right$_{H*\ H-H\%}$	purple$_{H*}$	egg$_{!H*}$
12	Initial N	To the left of THAT$_{L + H*\ H-L\%}$	silver$_{H*}$	stocking$_{H*}$
13	Mismatch Adj	And next$_{H*\ L-H\%}$	ORANGE$_{L + H*}$	stocking$_{noacc}$
14	Filler	Moving to the left$_{H*\ H-H\%}$	blue$_{H*}$	CANDY$_{L + H*}$
15	Initial Adj	FollowingTHAT$_{L + H*\ H-L\%}$	grey$_{H*}$	bell$_{!H*}$
16	Match Neither	After that$_{H*\ L-H\%}$	orange$_{H*}$	candy$_{!H*}$
17	Filler	Finally$_{H*\ H-H\%}$	white$_{H*}$	snowman$_{!H*}$
18	Filler	FIRST$_{L + H*\ H-L\%}$	yellow$_{H*}$	star$_{!H*}$
19	Mismatch DM	And NEXT$_{L + H*\ L-H\%}$	blue$_{H*}$	egg$_{!H*}$
20	Match Both	And THEN$_{L + H*\ L-H\%}$	GOLD$_{L + H*}$	egg$_{noacc}$
21	Filler	Moving to the right$_{H*\ H-L\%}$	yellow$_{H*}$	snowman$_{!H*}$
22	Mismatch DM	After THAT$_{L + H*\ L-H\%}$	red$_{H*}$	onion$_{!H*}$
23	Initial N Cont	To the right of that$_{H*\ H-L\%}$	red$_{H*}$	ANGEL$_{L + H*}$
24	Filler	Moving to the right$_{H*\ H-H\%}$	purple$_{H*}$	light bulb$_{!H*}$
25	Initial Adj	Following that$_{H*\ H-L\%}$	green$_{H*}$	onion$_{!H*}$
26	Match Neither	And then$_{H*\ L-H\%}$	grey$_{H*}$	candy$_{!H*}$
Tree 4				
1	Filler	At the very top$_{H*\ H-H\%}$	white$_{H*}$	tree$_{!H*}$
2	Both Initial	Starting on the LEFT$_{L + H*\ L-H\%}$	blue$_{H*}$	egg$_{!H*}$
3	Filler	To the right of that$_{H*\ H-H\%}$	yellow$_{H*}$	light bulb$_{!H*}$
4	Both Initial	On the right$_{H*\ H-H\%}$	brown$_{H*}$	bell$_{!H*}$
5	Filler	Starting on the left$_{H*\ H-H\%}$	purple$_{H*}$	angel$_{!H*}$
6	Initial N	Moving to the RIGHT$_{L + H*\ H-L\%}$	blue$_{H*}$	drum$_{H*}$
7	Match Both	And THEN$_{L + H*\ L-H\%}$	RED$_{L + H*}$	drum$_{noacc}$
8	Filler	Following that$_{H*\ H-H\%}$	purple$_{H*}$	tree$_{!H*}$
9	Initial N	To the right of THAT$_{L + H*\ H-L\%}$	orange$_{H*}$	angel$_{!H*}$
10	Match Neither	After that$_{H*\ L-H\%}$	red$_{H*}$	bell$_{!H*}$
11	Filler	Again on the LEFT$_{L + H*\ L-H\%}$	white$_{H*}$	stocking$_{!H*}$
12	Filler	To the right of that$_{H*\ H-H\%}$	yellow$_{H*}$	STAR$_{L + H*}$
13	Initial Adj	Following that$_{H*\ H-L\%}$	green$_{H*}$	bell$_{!H*}$
14	Mismatch Adj	And next$_{H*\ L-H\%}$	GOLD$_{L + H*}$	Dell$_{noacc}$
15	Mismatch DM	After THAT$_{L + H*\ L-H\%}$	brown$_{H*}$	egg$_{!H*}$
16	Initial N Cont	To the right of THAT$_{L + H*\ H-L\%}$	brown$_{H*}$	ONION$_{!H*}$
17	Match Neither	And then$_{H*\ L-H\%}$	blue$_{H*}$	stocking$_{!H*}$
18	Filler	First$_{H*\ H-H\%}$	purple$_{H*}$	LIGHT BULB$_{L + H*}$
19	Filler	Moving to the right$_{H*\ H-L\%}$	white$_{H*}$	star$_{!H*}$
20	Initial N	Following that$_{H*\ H-L\%}$	green$_{H*}$	candy$_{H*}$
21	Match Both	After NEXT$_{L + H*\ L-H\%}$	SILVER$_{L + H*}$	candy$_{noacc}$
22	Filler	To the right of that$_{H*\ L-H\%}$	yellow$_{H*}$	tree$_{!H*}$
23	Mismatch DM	And THEN$_{L + H*\ L-H\%}$	green$_{H*}$	angel$_{!H*}$
24	Initial N Cont	And next$_{H*\ H-L\%}$	green$_{H*}$	BALL$_{L + H*}$
25	Mismatch Adj	After that$_{H*\ L-H\%}$	GREY$_{H*}$	ball$_{noacc}$
26	Filler	FINALLY$_{L + H*\ H-L\%}$	white$_{H*}$	light bulb$_{!H*}$

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1988). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.

Bard, E.G.,& Aylett, M.P. (1999). The dissociation of de-accenting, giveness, and syntactic role in spontaneous speech. In *Proceedings of the Fourteenth International Conference of Phonetic Sciences*, San Francisco, 1753–1756.

Bard, E.G., Sotillo, C., Anderson, A.H., Doherty-Sneddon, G., & Newlands, A. (1995). The control of intelligibility in

running speech. In *Proceedings of the 13th International Congress of Phonetic Sciences*, Stockholm.

Bartels, C., & Kingston, J. (1994). Salient pitch cues in the perception of contrastive focus, In P. Bosch & R. van der Sandt (Eds.), *Focus and natural language processing*. IBM Working Papers on Logic and Linguistics, Vol. 6, (pp. 1–10), Heidelberg.

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes, 11*, 17–67.

Beckman, M.E., & Ayers, G M. (1997). *Guidelines for ToBI labelling*, vers 3.0 [manuscript]: Ohio State University.

Beckman, M.E., & Pierrehumbert, J.B. (1986). *Intonational structure in Japanese and English*, Phonology Yearbook, Vol. 3, 255–309.

Birch, S., & Clifton, C. (1995). Focus, accent, and argument structure: Effects on language comprehension. *Language and Speech, 38*, 365–391.

Bock, J. K., & Mazzella, J. R. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory and Cognition, 11*, 64–76.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the aharmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences, 17*, 97–110. University of Amsterdam. Can be downloaded as a pdf from http://fon/hum/uva.n1/paul/.

Boersma, P.W., & Weenick, D. (1992–2004). *Praat: doing phonetics by computer* (Version 4.2.17).

Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language, 37*, 83–96.

Bolinger, D. (1986). *Intonation and its parts*. London: Edward Arnold.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Geleerup: Lund.

Chafe, W. (1974). Language and consciousness. *Language, 50*, 111–133.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In C. Li (Ed.), *Subject and topic* (pp. 25–56). New York: Academic Press.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language, 47*, 30–49.

Cruttenden, A. (1986). *Intonation*. Avon: Cambridge University Press.

Cutler, A., Dahan, D., & Donselaar, W. v. (1997). Prosody in comprehension of spoken language: A literature review. *Language and Speech, 40*, 141–201.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course frequency effects in spoken-word recognition: Evidence from eye-movements. *Cognitive Psychology, 42*, 317–367.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16*, 507–534.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language, 47*, 292–314.

Davidson, D. (2001). Association with Focus in Denials. PhD dissertation, Michigan State Univerisity.

D'Imperio, M. (2000). The role of perception in defining tonal targets and their alignment. PhD dissertation, Ohio State University.

Face, T. (2001). Intonational marking of contrastive focus in Madrid Spanish. PhD Dissertation. Ohio State University.

Fery, C. (1993). *German intonational patterns*. Niemeyer.

Fischer, B. (1992). Saccadic reaction time: implications for reading, dyslexia and visual cognition. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 31–45). New York: Springer-Verlag.

Goldsmith, J. (1976). Autosegmental phonology, PhD dissertation. MIT, Cambridge, MA.

Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.

Grice, M., Baumann, S., & Benzmüller, R. (2005). German intonation in autosegmental–metrical phonology. In S.-A. Jun (Ed.), *Prosodic typology and transcription: A unified approach* (pp. 55–83). Oxford: OUP Press.

Grice, M., D'Imperio, M., Savino, M., & Avesani, C. (2005). Strategies for intonation labeling varieties of Italian. In S.-A. Jun (Ed.), *Prosodic typology and transcription: A unified approach* (pp. 362–389). Oxford: Oxford University Press.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English, part2. *Journal of Linguistics, 3*, 199–244.

Henderson, J. M., & Ferreira, F. (2004). Scene perceptions for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Hirschberg, J. (1993). Pitch accent in context: Predicting intonational prominence from text. *Artificial Intelligence, 63*, 305–340.

Ito, K. (2002). The interaction of focus and lexical pitch accent in speech production and dialogue comprehension: Evidence from Japanese and Basque. PhD dissertation. University of Illinois at Urbana-Champaign.

Ito, K., Speer, S.R., & Beckman, M. (2003). The influence of given-new status and lexical accent on intonation in Japanese spontaneous speech. Presentation to the Annual CUNY Conference on Sentence Processing, Boston, MA.

Ito, K., & Speer, S. R. (2006). Using interactive tasks to elicit natural dialogue. In P. Augurzky & D. Lenertova (Eds.), *Methods in empirical prosody research* (pp. 229–257). Mouton de Gruyter.

Kohler, K. (2005). Timing and communicative functions of pitch contours. *Phonetica, 62*, 88–105.

Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication, 34*, 391–405.

Ladd, R. D. (1996). *Intonational phonology* (Vol. 79). Cambridge: Cambridge University Press.

Ladd, R. D., & Morton, R. (1997). The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics, 25*, 313–342.

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics, 31*, 81–112.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information processing time with and without saccades. *Perception & Psychophysics, 53*, 372–380.

Nakatani, C. H. (1993). Accenting on pronouns and proper names in spontaneous narrative. Working Papers Department of Linguistics and Phonetics, 41, Lund, Sweden, 164–167.

Nakatani, C.H. (1997). The computational processing of intonational prominence: A functional prosody perspective. Ph.D. dissertation. Harvard University.

Needham, W. (1990). Semantic structure, information structure, and intonation in discourse production. *Journal of Memory and Language, 29*, 455–468.

Nooteboom, S. G., & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *Journal of Acoustical Society of America, 82*, 1512–1524.

Pierrehumbert, J.B. (1980). The phonology and phonetics of English intonation. PhD dissertation. Massachusetts Institute of Technology.

Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. Cambridge: MIT Press.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 342–365). Cambridge: MIT Press.

Prieto, P., van Santen, J., & Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics, 23*, 429–451.

Raaihmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with ''The language-as-fixed-effect-fallacy'': Common misconceptions and alternative solutions. *Journal of Memory and Language, 41*, 416–426.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of Optical Society of America, 57*, 1030–1033.

Schafer, A. J., Speer, S. R., & Warren, P. (2003). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In Tanenhaus & Trueswell (Eds.), *World situated language use:Psycholinguistic linguistic, and computational perspectives on bridging the product and action tradition* (pp. 209–225). Cambridge: MIT Press.

Sedivy, J., Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Carlson, G. (1995). Using intonationally-marked presuppositional information in on-line language processing: Evidence from eye movements to a visual model. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, Hillsdale, NJ.

Sedivy, J., Tanenhaus, M., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109–147.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language, 48*, 103–130.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Effects of visual context in the resolution of temporary syntactic ambiguities in spoken language comprehension. *Cognitive Psychology, 45*, 447–481.

Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K. M., & Sedivy, J. C. (1996). Using eye-movements to study spoken language comprehension: Evidence for visually-mediated incremental interpretation. In T. Inui & J. McClelland (Eds.), *Attention & Performance XVI: Integration in Perception and Communication* (pp. 457–478). Cambridge: MIT Press.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In *Handbook in perception and cognition*. In Eimas & Miller (Eds.). *Speech language and communication* (Vol. 11, pp. 217–262). NY: Academic Press.

Terken, J. M. B., & Hirschberg, J. (1994). Deaccentuation of words representing ''given'' information: Effects of persistence of grammatical function and surface position. *Language and Speech, 37*, 125–145.

Terken, J. M. B., & Nooteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes, 2*, 145–163.

Trueswell, J. C., & Tanenhaus, M. K. (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*. Cambridge, MA: MIT Press.

Venditti, J. J. (1997). Japanese ToBI labelling guidelines. *Ohio State University Working Papers in Linguistics, 50*, 62–72.

Venditti, J.J. (2005). The J-ToBI model of Japanese intonation. In Jun, S.-A. (Ed.) *Prosodic Typology and Transcription: A Unified Approach*, (pp. 172-200). Oxford: Oxford University Press. Online at http://ling.osu.edu/research/phonetics/J_ToBI/.

Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual, and motor control aspects. In K. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 353–393). Amsterdam: Elsevier.